

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:03:55 ; Search time 13.4328 Seconds
(without alignments)
198.395 Million cell updates/sec

Title: US-10-036-371-2

Perfect score: 107

Sequence: 1 IVGGYTCQNTVPVQVLSNS 20

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	20	1	AA03086
2	107	100.0	20	21	N-terminus of Bovine
3	107	100.0	22	20	Amino terminal seq
4	107	100.0	22	26	Human trypsinogen
5	107	100.0	223	21	Trypsinogen B
6	107	100.0	229	21	Bovine trypsinogen
7	107	100.0	229	21	Bovine trypsinogen
8	107	100.0	230	15	Bovine trypsinogen
9	107	100.0	231	21	Recombinant trypsin
10	107	100.0	231	22	Bovine met-phe-tryp
					Trypsinogen analog

11	98	91.6	223	23	AB083322	Partial trypsin se
12	96	89.7	247	18	AA08475	Porcine trypsinoge
13	92	86.0	20	21	AAV78977	Canine anionic try
14	92	86.0	246	21	AAV78975	Canine anionic try
15	87	81.3	240	19	AAW57740	Trypsinogen-like p
16	87	81.3	247	9	AA081243	Human spleen tryps
17	87	81.3	247	16	AA082703	Human pancreatic tr
18	87	81.3	247	21	AAV78974	Canine cationic tr
19	86	80.4	247	21	AA035701	Human trypsin hL a
20	85	79.4	224	20	AA031160	Human trypsin seri
21	85	79.4	225	22	AA095503	Human trypsinogen
22	85	79.4	230	20	AAW93488	Human trypsinogen
23	85	79.4	241	21	AA021348	Human amyloid beta
24	85	79.4	246	19	AA064260	Human trypsinogen
25	85	79.4	254	21	AA074758	Human pancreatic tr
26	85	79.4	254	21	AA074758	Human pancreatic tr
27	84	78.5	247	23	AAU87693	Atlantic cod hydro
28	67	62.6	25	20	AAV33346	Salmon enzyme i mu
29	67	62.6	25	21	AAV93936	N-terminal of an a
30	67	62.6	25	22	AA021357	N-terminal of slam
31	67	62.6	25	23	AA021357	Salmon enzyme i mu
32	67	62.6	240	23	AB04644	Engraulis japonicu
33	67	62.6	37	21	AAV93946	N-terminal of an a
34	65	60.7	37	23	AA021360	Atlantic cod multi
35	65	60.7	241	23	AB04645	Engraulis japonicu
36	65	60.7	225	22	AA031579	Amino acid sequenc
37	61	57.0	23	21	AAV93937	N-terminal of an a
38	60	56.1	23	21	AAV93937	N-terminal of acta
39	60	56.1	23	22	AA057943	N-terminal of acta
40	60	56.1	23	22	AA057944	Atlantic cod hydro
41	60	56.1	25	20	AAV33347	Arilla-derived mult
42	60	56.1	25	20	AAV33348	Atlantic cod multi
43	60	56.1	25	21	AA021358	Atlantic cod multi
44	60	56.1	25	21	AA021358	Atlantic cod multi
45	60	56.1	25	23	AA021358	Atlantic cod i mu

ALIGNMENTS

RESULT 1
AA03086
ID AA03086 standard; peptide; 20 AA.
XX
AC AA03086;
XX
DT 10-OCT-2000 (first entry)
XX
XZ N-terminus of bovine trypsin.
XX
XZ Trypsin; bovine; serine protease; Trichoderma; coagulant; hypertensive;
KW antinflammatory; leather preparation; silk treatment.
XX
XX Bos taurus.
XX
PN JP2000116377-A.
XX
PD 25-APR-2000.
XX
PF 08-OCT-1998; 98JP-0303263.
XX
PR 08-OCT-1998; 98JP-0303263.
XX
(AMAN) AMANO PHARM KK.
XX
XZ WPI; 2000-369402/32.
XX
PT A new serine protease and its preparation, used clinically in blood
XX coagulation, hypotension and anti-inflammation .
XX
PS Example 3; Page 7; 9pp; Japanese.
XX
CC The invention relates to a novel serine protease from Trichoderma sp.

CC No. 9064. The N-terminus of this protease is given in AAB03085. The novel
 CC protease has trypsin-like activity, specifically cleaving peptide chains
 CC on the carboxyl side of a basic amino acid (e.g., arginine or lysine).
 CC It has a pH optimum of 7-8, and is stable between pH 5 and pH10. It has
 CC a temperature optimum of approximately 40 degrees Celsius, and is stable
 CC between 40 and 50 degrees Celsius. The invention also relates to a
 CC method for the preparation of the novel serine protease, and the use of
 CC the protease in protein degradation. The enzyme has coagulant,
 CC hyperinvasive and anti-inflammatory effects. It may also be used in the
 CC preparation of leather, for raw silk treatment and for the preparation of
 CC protein hydrolysate. Sequences AAB03086-B03089 represent the N-terminal
 CC 20 amino acids of trypsin from a variety of organisms.

XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 107; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
 Db 1 IVGGYTCGANTVPYQVSLNS 20
 |||||
 |||||

RESULT 2
 ID AAB31574 standard; peptide; 20 AA.

XX AAB31574;
 XX AAB31574;
 XX 20-APR-2001 (first entry)

XX Amino terminal sequence of bovine trypsins.

XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;
 XX arthritis; inflamed joints; bursitis; osteoarthritis; septic arthritis;
 XX rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
 XX systemic lupus erythematosus; psoriasis; eczema; hemorrhoid; proctitis;
 XX acne; eczema; facial seborrheic eczema; seborrheic infection; proctitis;
 XX athlete's foot; fistulae infection; ulcer; navel infection; wrinkle;
 XX scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;
 XX fungal infection; autoimmune disease.

XX Bos sp.

XX WO2000078332-A2.

XX 28-DEC-2000.

XX 15-JUN-2000; 2000WO-IS00005.

XX 18-JUN-1999; 99IS-0005086.

XX (BJAR/) BJARNASON J B.

XX Bjarnason JB;

XX WPI; 2001-091493/10.

XX Fish serine proteinase, useful as a cosmetic, medicament for treating
 XX eczema, psoriasis, arthritis, and in the manufacture of the medicament
 XX for treating, preventing pathogenic diseases involving receptor
 XX mediated binding -

XX Disclosure; Page 5; 38pp; English.

XX The specification describes a fish serine proteinase. The proteinases
 XX are useful as medicaments, for treating and preventing a disease in a
 XX human or an animal such as pain, acute inflammation, osteoarthritis,
 XX inflammation, arthritis, inflamed joints, bursitis, eczema, psoriasis,
 XX rheumatoid arthritis, juvenile rheumatoid arthritis, septic arthritis,
 XX fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, rash,
 XX psoriasis, acne, eczema, facial seborrheic eczema, eczema of the hands,

CC face or neck, foreskin infections, athlete's foot, fistulae infections,
 CC infected topical ulcers, navel infections in newborns, wrinkles, scars,
 CC keloids, boils, warts and allergic itch, hemorrhoids, wound
 CC infections, wounds from burns, fungal infections and immunological and
 CC autoimmune diseases. They are also useful for removing dead or peeling
 CC skin from otherwise healthy skin, and for treating or preventing a
 CC disease in which pathogenesis is caused by bacteria, virus, fungus,
 CC parasite or a protozoan or a receptor mediated binding is involved.
 CC The present sequence represents the amino terminal of bovine trypsins.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 107; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
 Db 1 IVGGYTCGANTVPYQVSLNS 20
 |||||
 |||||

RESULT 3
 ID AAB81767 standard; peptide; 223 AA.

XX AAB81767;
 XX AAB81767;
 XX 29-JAN-1999 (first entry)

XX Bovine TRYP peptide fragment.

XX DHFR; dihydrofolic acid reductase; protein function; trypsin; bovine;
 XX ribonuclease; myoglobin; database; homology; resemblance.

XX Bos taurus.

XX JPI0287696-A.

XX 27-OCT-1998.

XX 11-APR-1997; 97JP-0093577.

XX 11-APR-1997; 97JP-0093577.

XX (IYAK-) IYAKU BUNSHI SEKKEI KENKYUSHO KK.

XX WPI; 1999-018384/02.

XX Determination of protein biological function - comprises use of
 XX amino acid sequences database containing the relevant information

XX Example 1; Fig 1; 11pp; Japanese.

XX This sequence is used in the creation of a database containing the
 XX information for amino acid sequence of protein with at least 1
 XX biological function with added a score on importance of expression of
 XX the biological information for each amino acid residue. The database is
 XX useful for determination of unknown biological function of a protein or
 XX polypeptide based on the homology of amino acid sequence, e.g. steric
 XX structure of protein, and includes retrieval and evaluation of high
 XX homologous relationship for the determination of mostly resembling
 XX protein. The database allows for correct and rapid retrieval and
 XX presumption of protein and polypeptide having biological functions.

XX SQ Sequence 223 AA;

Query Match 100.0%; Score 107; DB 20; Length 223;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
 Db 1 IVGGYTCGANTVPYQVSLNS 20
 |||||
 |||||

PI Howard JA, Hood E;
 XX WPI, 2000-195105/17.
 DR N-PSDB; AA292585.
 XX
 PT Transgenic plant containing DNA encoding a heterologous protein, used
 for the commercial production of protease -
 XX
 XX Example 1; Page 28-29; 32pp; English.
 XX The invention relates to a transgenic plant that comprises DNA encoding
 an active or inactive form of a heterologous protease (e.g., trypsin)
 CC in the nucleotide sequence encoding the protease is operably linked to a
 CC promoter to effect expression of the protease in the plant. The
 CC transgenic plant is used in a method of producing protease in commercial
 CC quantities, comprising the extraction of the protease from the plant
 CC tissue. The source of trypsin is the organs of animals. One
 CC of the difficulties with this source is that there is considerable
 CC contamination by other proteases. Chymotrypsin is one of the additional
 CC in the contaminant, and may cleave the product in an undesired manner.
 CC The method of the invention overcomes the problem of contamination by
 CC other additional proteases by directing the expression of the protease
 CC to the seed of the plant. This problem of expressing the protease and
 CC having it digest the proteins in the plant, resulting in cell death is
 CC overcome. The present sequence addresses bovine trypsinogen. cDNA
 CC (Genbank accession #P00760) encoding this protease was used in the
 CC construction of a plant expression plasmid in the exemplifications of
 CC the invention.
 XX
 XX Query Match 100.0%; Score 107; DB 21; Length 229;
 XX Best Local Similarity 100.0%; Pred. No. 3e-08;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 IVGGYTCGANTVPYQVSLNS 20
 DB 7 IVGGYTCGANTVPYQVSLNS 26
 XX
 XX RESULT 7
 XX AAARS3638
 ID AAARS3638 standard; Protein; 230 AA.
 AC AAARS3638;
 DT 01-DEC-1994 (first entry)
 XX Bovine trypsinogen.
 XX Cattle; cow; trypsinogen; enzyme; protease; proinsulin; insulin;
 XX hormone; plasmid pRMG4.
 XX Bos taurus.
 XX EP597681-A.
 XX 18-MAY-1994.
 XX 10-NOV-1993; 93EP-0308959.
 XX 13-NOV-1992; 92US-0977703.
 XX (ELIL) LILLY & CO ELI.
 XX Greaney MG, Rostack PR;
 XX WPI, 1994-160671/20.
 XX Expression vectors for bovine trypsin and bovine trypsinogen -
 XX for cleavage of zymogens into active drugs, e.g. pro-insulin
 XX conversion into insulin

PS Disclosure; Page 27; 35pp; English.
 XX This protein may be expressed by E. coli/plasmid pRMG7 and is able
 CC to cleave zymogens into active drugs, e.g. pro-insulin conversion
 CC into insulin.
 XX
 XX Sequence 230 AA;
 Query Match 100.0%; Score 107; DB 15; Length 230;
 XX Best Local Similarity 100.0%; Pred. No. 3e-08;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 IVGGYTCGANTVPYQVSLNS 20
 DB 8 IVGGYTCGANTVPYQVSLNS 27
 XX
 XX RESULT 8
 XX AAAY91926
 ID AAAY91926 standard; Protein; 231 AA.
 AC AAAY91926;
 DT 19-JUL-2000 (first entry)
 XX Recombinant trypsin.
 XX DE Recombinant trypsin.
 XX KW Recombinant trypsin; trypsinogen analogue; mutated bovine trypsinogen;
 XX leader sequence; trypsin activity; recombinant protein production;
 XX inactive zymogen.
 XX Synthetic.
 XX Bos taurus.
 XX Key Location/Qualifiers
 XX Peptide 1..2
 XX Peptide 3..7 /note= "optionally absent, see disclosure"
 XX Peptide /label= leader_sequence
 XX Protein 8..231
 XX /label= mature_protein
 XX WO200017332-A1.
 XX 30-MAR-2000.
 XX 15-SEP-1999; 99WO-US21047.
 XX 21-SEP-1998; 98US-0101213.
 XX (ELIL) LILLY & CO ELI.
 XX Hanquier JM, Hershenberger CL, Deepiancq D, Larson JL, Rostack PR;
 XX WPI, 2000-283565/24.
 XX N-PSDB; AAA08526.
 XX New trypsinogen analog useful for the production of recombinant trypsin
 XX has a modified leader sequence not cleavable by trypsin or trypsin-like
 XX enzymes
 XX Claim 22; Page 51-52; 56pp; English.
 XX This sequence shows a claimed recombinant trypsin. The trypsin is
 CC produced by cleavage of a trypsinogen analogue (AAAY91925). A wild type
 CC bovine trypsinogen was mutated to destroy the trypsin cleavage site.
 CC The lys residue present in the leader sequence of the native bovine
 CC trypsinogen protein was mutated to an Asp residue. The vector was
 CC constructed such that DNA encoding a (Glu-Ala)₂ peptide was inserted
 CC between the C-terminus of the alpha factor signal and the N-terminus of
 CC the trypsinogen analogue leader sequence to improve the yield of the
 CC secreted protein. The specification claims an isolated trypsinogen
 CC analogue comprising a protein having trypsin activity and a leader

CC sequence having at least two amino acids which are not Lys or Arg.
 CC The trypsin derived from the recombinant trypsinogen is useful for the
 CC characterization of other proteins, and in the manufacture of other
 CC recombinant bioproducts, for example to cleave leader sequences of proteins
 CC from small recombinant proteins expressed initially in inactive zymogen form.
 CC The present method properly folded yet is not activated until after
 CC purification from fermentation broth or cell extracts. This is
 CC accomplished through the expression of a single chain trypsinogen
 CC analogue where the leader sequence is modified such that it lacks a
 CC trypsin-like enzyme cleavage site. Specifically the trypsinogen
 CC analogues of the present invention lack a lysin or arginine in the
 CC N-terminal leader sequence of the protein to prevent auto-activation or
 CC activation by endogenous host cell enzymes.

XX Sequence 231 AA;
 Query Match 100.0%; Score 107; DB 21; Length 231;
 Best Local Similarity 100.0%; Pred. No. 3e-08; 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0;

Oy 1 IVGGYTCGANTPYQVSLNS 20
 Db 9 IVGGYTCGANTPYQVSLNS 28

RESULT 9
 AAB80953
 ID AAB80953 standard; Protein; 231 AA.

XX AAB80953;
 DT 06-JUN-2001 (first entry)
 DE Bovine met-phe-trypsinogen.
 XX Trypsinogen; bovine; trypsin; serine protease.
 XX Bos sp.

XX Key Location/Qualifiers
 FT Peptide 4..8
 FT /note= "Leader peptide, cleaved to produce trypsin"

XX W0200119970-A2.

XX 22-MAR-2001.

XX 05-SEP-2000; 200MO-US20813.

XX 15-SEP-1999; 99MO-0154019.

XX (ELIL) LILLY & CO ELI.

XX Hanquier JM, Hersberger CL, Larson JL, Rostack PR;

XX WPI; 2001-273425/28.

XX N-PSDB; AAF81479.

XX New chymotrypsin-free trypsin and trypsinogen useful for manufacturing
 PT recombinant protein pharmaceuticals and pure trypsin .

XX Claim 20; Fig 1; 55pp; English.

XX The present sequence is bovine met-phe-trypsinogen. Trypsin is a serine
 CC protease which cleaves the peptide bond on the carboxy-terminus of basic
 CC amino acid residues. Trypsin is synthesised in a slightly longer
 CC catalytically inactive form: trypsinogen, which is a leader
 CC sequence removed to trypsin. The amino acid sequence of the present
 CC protein is the same as trypsinogen, and is present at the amino-terminus.
 CC This trypsin has two additional residues at the amino terminus: Met and
 CC Phe. Bovine met-phe-trypsinogen is useful for the manufacture of
 CC recombinant protein pharmaceuticals. High purity trypsin products are

CC produced by expressing trypsinogen inside a prokaryotic cell which is
 CC then isolated and activated to form trypsin.

XX Sequence 231 AA;

Query Match 100.0%; Score 107; DB 22; Length 231;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTPYQVSLNS 20
 Db 9 IVGGYTCGANTPYQVSLNS 28

RESULT 10

AA919225
 ID AA919225 standard; Protein; 233 AA.

XX AA919225;

DT 19-JUL-2000 (first entry)

XX Trypsinogen analogue.

XX Trypsinogen analogue; mutated bovine trypsinogen; leader sequence;
 KW trypsin activity; recombinant protein production; inactive zymogen.

XX Synthetic.

OS Bos taurus.

XX Key Location/Qualifiers

FT Peptide 1..55= "optionally absent, see disclosure"

FT Peptide 1..4

FT /label= linker_peptide

FT /note= "increases yield of secreted protein"

FT Peptide 5..10

FT /label= leader_sequence

FT Misc-difference 10

FT /note= "native bovine trypsinogen protein was

FT mutated to an Asp residue"

FT Protein 11..233

FT /label= mature_protein

FT /note= "trypsin"

FT Modified-eite 169

FT /note= "potential N-glycosylation"

XX W0200017332-A1.

XX 30-MAR-2000.

XX 15-SEP-1999; 99MO-US21047.

XX 21-SEP-1998; 98US-0101213.

XX (ELIL) LILLY & CO ELI.

XX Hanquier JM, Hersberger CL, Desplancq D, Larson JL, Rostack PR;

XX WPI; 2000-283565/24.

XX N-PSDB; AAA08525.

XX New trypsinogen analog useful for the production of recombinant trypsin
 PT has a modified leader sequence not cleavable by trypsin or trypsin-like
 PT enzymes

XX Claim 5, 6; Page 47-48; 56pp; English.

XX This sequence shows a trypsinogen analogue. The wild type bovine
 CC trypsinogen was mutated to destroy the trypsin cleavage site. The lys
 CC residue present in the leader sequence of the native bovine trypsinogen
 CC protein was mutated to an Asp residue. The vector was constructed such
 CC that DNA encoding a (Glu-Ala)2 peptide was inserted between the


```

Db      25 IVGGYTCANSPVQVSLNS 44
      ||||| |||:|||||
RESULT 13
ID      AAY78977 standard; protein; 20 AA.
AC      AAY78977;
XX
XX
DT      05-JUN-2000 (first entry)
XX
XX      Canine anionic trypsinogen immunogenetic fragment #2 amino acid sequence.
DE
XX      Anionic trypsinogen; dog; monoclonal antibody production; detect;
XX      diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
XX      extrapancreatic hyposcretion; immunogenic fragment.
KW
XX      Canis familiaris.
OS
XX      MPI; 2000-009739-A1.
XX
XX      24-FEB-2000.
XX
XX      09-AUG-1999; 99WO-JP04299.
XX
XX      10-AUG-1998; 98JP-0236609.
XX
XX      10-MAR-1999; 99JP-0063990.
XX
XX      (FUJY ) FUJI YAKUHIN KOGYO KK.
XX
XX      Maritani T, Ashida Y, Yamada T;
XX
XX      MPI; 2000-206018/18.
XX
XX      Anti-canine trypsin monoclonal antibody, useful as reagent for quick
XX      and accurate detection and quantitation of trypsin and/or trypsin-like
XX      immunoreactants in various forms in diagnosis e.g. of pancreatic
XX      diseases -
XX
XX      Example 2; Page 33; 67pp; Japanese.
XX
XX      This sequence represents an immunogenic fragment of the canine anionic
XX      trypsinogen amino acid sequence. The fragment corresponds to amino acid
XX      residues 24 to 43 of the anionic trypsinogen protein sequence
XX      (see AAY78975). The invention relates to monoclonal antibodies with
XX      specificity against canine trypsin, or canine trypsin-related
XX      substances. The antibodies are highly specific and can be used as reagent
XX      for quick and accurate detection and quantitation of canine trypsin and
XX      canine trypsin-like immunoreactants in various forms. The antibodies can
XX      be used in the diagnosis of diseases such as pancreatitis, pancreatic
XX      cancer, renal insufficiency and extrapancreatic hyposcretion.
XX
XX      Sequence 20 AA;
SQ
XX
XX      Query Watch      86.0%; Score 92; DB 21; Length 20;
XX      Best Local Similarity 85.0%; Pred. No. 3/6e-07;
XX      Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX      QY      1 IVGGYTCGANTVPVQVSLNS 20
XX      ||||| |||:|||||
XX      DB      1 IVGGYTCRNSVPVQVSLNS 20
XX
XX
XX      RESULT 14
XX      AAY78975
XX      AAY78975 standard; protein; 246 AA.
XX
XX      AAY78975;
XX
XX      05-JUN-2000 (first entry)
XX
XX      Canine anionic trypsinogen amino acid sequence.
DE

```

```

XX      Anionic trypsinogen; dog; monoclonal antibody production; detect;
XX      diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
XX      extrapancreatic hyposcretion.
XX
XX      Canis familiaris.
OS
XX      MPI; 2000-009739-A1.
XX
XX      24-FEB-2000.
XX
XX      09-AUG-1999; 99WO-JP04299.
XX
XX      10-AUG-1998; 98JP-0236609.
XX
XX      10-MAR-1999; 99JP-0063990.
XX
XX      (FUJY ) FUJI YAKUHIN KOGYO KK.
XX
XX      Maritani T, Ashida Y, Yamada T;
XX
XX      MPI; 2000-206018/18.
XX
XX      Anti-canine trypsin monoclonal antibody, useful as reagent for quick
XX      and accurate detection and quantitation of trypsin and/or trypsin-like
XX      immunoreactants in various forms in diagnosis e.g. of pancreatic
XX      diseases -
XX
XX      Claim 3; Page 64-65; 67pp; Japanese.
XX
XX      This sequence represents the canine anionic trypsinogen amino acid
XX      sequence. The protein was isolated from the canine pancreas. The
XX      invention relates to monoclonal antibodies with specificity against
XX      canine trypsin, or canine trypsin-related substances. The antibodies are
XX      highly specific and can be used as reagent for quick and accurate
XX      diagnosis of diseases such as pancreatitis, pancreatic cancer, renal
XX      insufficiency and extrapancreatic hyposcretion.
XX
XX      Sequence 246 AA;
SQ
XX
XX      Query Match      86.0%; Score 92; DB 21; Length 246;
XX      Best Local Similarity 85.0%; Pred. No. 6e-06;
XX      Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX      QY      1 IVGGYTCGANTVPVQVSLNS 20
XX      ||||| |||:|||||
XX      DB      24 IVGGYTCRNSVPVQVSLNS 43
XX
XX      RESULT 15
XX      AAY57740
XX      AAY57740 standard; protein; 240 AA.
XX
XX      AAY57740;
XX
XX      16-SEP-1998 (first entry)
XX
XX      Trypsinogen-like protein.
XX
XX      Trypsinogen-like protein; human.
XX
XX      Homo sapiens.
XX
XX      JP10090980-A.
XX
XX      21-APR-1998.
XX
XX      26-SEP-1996; 96JP-0273923.
XX
XX      26-SEP-1996; 96JP-0273923.
XX
XX      (SHIS ) SHISEIDO CO LTD.
PA

```

XX WPI; 1996-289873/26.
DR N-PSDB; AAV24548.
XX DNA coding for trypsinogen-like protein
PT
XX
XX Claim 3; Page 4-5; 7pp; Japanese.
PS
XX
XX This sequence represents a human trypsinogen-like protein, and is encoded
CC by the DNA of the invention.
XX
XX SQ Sequence 240 AA;
Query Match 81.3%; Score 87; DB 19; Length 240;
Best Local Similarity 80.0%; Pred. NO. 3.3e-05;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 IVGGVTCGANTVPVQVSLNS 20
DB 17 IVGGITCENSIPVQVSLNS 36
Search completed: February 12, 2003, 10:22:22
Job time : 15.4328 secs

. . . .


```

; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Atlantic cod
US-09-938-269-16

Query Match          94.7%; Score 90; DB 9; Length 25;
Best Local Similarity 85.0%; Pred. No. 9.8e-09;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXKHQAHOVSLS 20
Db 1 IVGGYCTHQAHOVSLS 20

RESULT 3
US-09-938-269-17
; Sequence 17: Application US/09938269
; Publication No. US20030007951A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; APPLICANT: Cowling, Didier S.P.
; APPLICANT: Hubbel, Jeffrey A.
; APPLICANT: van de Wetering, Petra
; TITLE OF INVENTION: Treatment of Trauma
; FILE REFERENCE: 314572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Atlantic cod
US-09-938-269-17

Query Match          94.7%; Score 90; DB 9; Length 37;
Best Local Similarity 85.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXKHQAHOVSLS 20
Db 1 IVGGYCEHQAHOVSLS 20

RESULT 4
US-10-036-371-8
; Sequence 8: Application US/10036371
; Publication No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: 81691-284960 COSMETIC USE
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (9)
; OTHER INFORMATION: K or R

Query Match          94.7%; Score 90; DB 12; Length 225;
Best Local Similarity 90.0%; Pred. No. 9.3e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVGGYCXKHQAHOVSLS 20
Db 1 IVGGYCTHQAHOVSLS 20

RESULT 5
US-10-036-371-1
; Sequence 1: Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: 81691-284960 COSMETIC USE
; CURRENT APPLICATION NUMBER: US/10/036,371

```



```

; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBERS OF SEQ ID NOS: 8
; OTHER INFO: Patent in Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (6)
; OTHER INFORMATION: Q or E
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (8)
; OTHER INFORMATION: E or T
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (9)
; OTHER INFORMATION: K or R
; US-10-036-371-1

Query Match 93.7%; Score 89; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQHQVSLNS 20
Db 1 IVGGYCXHXSHQHQVSLNS 20

RESULT 6
US-09-938-269-14
; Sequence 14, Application US/09938269
; Publication No. US20030007951A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; INVENTOR: Franklin, Richard L.
; APPLICANT: Subbaiah, Jeffrey A.
; INVENTOR: Subbaiah, Jeffrey A.
; APPLICANT: van de Wetering, Petra
; INVENTOR: van de Wetering, Petra
; TITLE OF INVENTION: Treatment of Trauma
; FILE REFERENCE: 314572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Salmon enzyme
; US-09-938-269-14

Query Match 82.1%; Score 78; DB 9; Length 25;
Best Local Similarity 75.0%; Pred. No. 9.2e-07;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQHQVSLNS 20
Db 1 IVGGYCKKAYCAQVQVSLNS 20

RESULT 7
US-10-021-368-10
; Sequence 10, Application US/10021368
; Patent No. US20020106367A1
; GENERAL INFORMATION:
; APPLICANT: Band, Vmla
; INVENTOR: Band, Vmla
; TITLE OF INVENTION: RES-1 POLYPEPTIDES, DNA, AND RELATED
; MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/021,368
; FILING DATE: 12-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATIONS:
; APPLICATION NUMBER: 09/201,038
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-021-368-10

Query Match 69.5%; Score 66; DB 12; Length 271;
Best Local Similarity 53.5%; Pred. No. 0.00099;
Matches 15; Conservative 1; Mismatches 4; Indels 8; Gaps 1;

Qy 1 IVGGYX-----CXHXSHQHQVSLNS 20
Db 10 IVGGYSSTRYPIECKYASQHQVSLNS 37

RESULT 8
US-09-923-779-154
; Sequence 154, Application US/09923779
; Patent No. US2002007621A1
; GENERAL INFORMATION:
; APPLICANT: Fyle, Ruth A.
; INVENTOR: Fyle, Ruth A.
; APPLICANT: Kales, Michael D.
; INVENTOR: Kales, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF PANCREATIC CANCER
; FILE REFERENCE: 210121-553
; CURRENT APPLICATION NUMBER: US/09/923,779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-923-779-154

Query Match 67.4%; Score 64; DB 10; Length 247;
Best Local Similarity 65.0%; Pred. No. 0.0019;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQHQVSLNS 20

```

Db 24 IVGGYCENSVPYQVSLNS 43

RESULT 9
US-10-036-371-2
; Sequence 2, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Bovine sp.
US-10-036-371-2

Query Match 66.3%; Score 63; DB 12; Length 20;
Best Local Similarity 60.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYKXCHXSQAHQVSLNS 20
||||| : : : :
Db 1 IVGGYTCGANTVPYQVSLNS 20
||||| : : : :
RESULT 10
US-09-910-071-14
; Sequence 14, Application US/09910071
; Patent No. US20020116146A1
; GENERAL INFORMATION:
; APPLICANT: Tomikawa, Mayumi
; APPLICANT: Akawa, Seichi
; APPLICANT: Matsumura, Fumiko
; TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Simil
; TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three
; FILE REFERENCE: 522.192102
; CURRENT APPLICATION NUMBER: US/09/910,071
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/014,667
; PRIOR FILING DATE: 1993-02-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 223
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: amino acid sequence of trypsin
US-09-910-071-14

Query Match 65.3%; Score 62; DB 10; Length 223;
Best Local Similarity 60.0%; Pred. No. 0.0037;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYKXCHXSQAHQVSLNS 20
||||| : : : :
Db 1 IVGGYTCGANTVPYQVSLNS 20
||||| : : : :
RESULT 11
US-09-925-300-1093
; Sequence 1093, Application US/09925300

; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-18
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1093
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1093

Query Match 57.9%; Score 55; DB 10; Length 235;
Best Local Similarity 55.0%; Pred. No. 0.055;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 IVGGYKXCHXSQAHQVSLNS 20
||||| : : : :
Db 87 IVGGMECEKHSQPQWAVYS 106
||||| : : : :
RESULT 12
US-10-021-368-9
; Sequence 9, Application US/10021368
; Patent No. US20020106367A1
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 12-Dec-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/201,038
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:

THIS PAGE BLANK (USPTO)

ADDRESSES: Dechert Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FastSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/385,540A
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/600,273
 FILING DATE: 08-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: 314572-101A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-385-540A-19

Query Match 94.7%; Score 90; DB 2; Length 25;
 Best Local Similarity 85.0%; Pred. No. 8.7e-10;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQAHQVSLNS 20
 Db 1 IVGGYECTKHSQAHQVSLNS 20

RESULT 3
 US-08-385-540A-20
 Sequence 20; Application US/08385540A
 Patent No. 5945102
 GENERAL INFORMATION:
 APPLICANT: de Faire, Johan
 TITLE OF INVENTION: Wound Care With Multifunctional
 TITLE OF INVENTION: Enzyme
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Dechert Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FastSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/385,540A
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/600,273
 FILING DATE: 08-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: 314572-101A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-385-540A-20
 Query Match 94.7%; Score 90; DB 2; Length 25;
 Best Local Similarity 85.0%; Pred. No. 8.7e-10;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQAHQVSLNS 20
 Db 1 IVGGYECTKHSQAHQVSLNS 20

RESULT 4
 US-08-600-273A-19
 Sequence 19; Application US/08600273A
 Patent No. 5958406
 GENERAL INFORMATION:
 APPLICANT: de Faire, Johan
 APPLICANT: Franklin, Richard L.
 APPLICANT: Kay, John
 TITLE OF INVENTION: Acne Treatment With Multifunctional
 TITLE OF INVENTION: Enzyme
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Dechert Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: FastSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/600,273A
 FILING DATE: 08-FEB-1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/486,820
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/385,540
 FILING DATE: 08-FEB-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: 314572-101C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-620-3214
 TELEFAX: 609-620-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-600-273A-19

Query Match 94.7%; Score 90; DB 2; Length 25;

```

Best Local Similarity 85.0%; Pred. No. 8.7e-10;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYXXHQAQVSLNS 20
Db 1 IVGGYCTHQAQVSLNS 20

RESULT 5
US-08-600-273A-20
; Sequence 20, Application US/08600273A
; Patent No. 5358406
; GENERAL INFORMATION:
; APPLICANT: de Faire, Johan
; APPLICANT: Franklin, Richard L.
; APPLICANT: Kay, John
; TITLE OF INVENTION: Acne Treatment With Multifunctional
; TITLE OF INVENTION: Enzyme
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,820
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/486,820
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 314572-101B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-620-3214
; TELEFAX: 609-620-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; QUERY MATCH
; Best Local Similarity 85.0%; Pred. No. 8.7e-10;
; Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYXXHQAQVSLNS 20
Db 1 IVGGYCTHQAQVSLNS 20

RESULT 7
US-08-486-820-20
; Sequence 20, Application US/08486820
; Patent No. 6030612
; GENERAL INFORMATION:
; APPLICANT: de Faire, Johan
; APPLICANT: Franklin, Richard L.
; APPLICANT: Kay, John
; TITLE OF INVENTION: Antimicrobial Uses Of Multifunctional
; TITLE OF INVENTION: Enzyme
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,820

```



```

; APPLICANT: de Faire, Johan
; APPLICANT: Franklin, Richard L.
; APPLICANT: Kay, John
; TITLE OF INVENTION: Antimicrobial Uses Of Multifunctional
; TITLE OF INVENTION: Enzyme
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ USA
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION NUMBER: US/08/486,820
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/486,820
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 314572-101B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-620-3214
; TELEFAX: 609-620-3259
;
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-486-820-18

Query Match      82.1%; Score 78; DB 3; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.2e-07;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYKXCHSQAHQVLSNS 20
Db 1 IVGGTCKATSAQVLSNS 20

RESULT 11
US-09-220-731-17
; Sequence 17, Application US/09220731A
; Patent No. 6232088
; GENERAL INFORMATION:
; APPLICANT: Phairen Medical, Inc.
; APPLICANT: Richard L. Franklin
; APPLICANT: Yves St. Pierre
; TITLE OF INVENTION: Treatment and Prevention of Immune
; FILE REFERENCE: 314572-101D
; CURRENT APPLICATION NUMBER: US/09/220,731A
; CURRENT FILING DATE: 1998-12-24
; EARLIER FILING DATE: 1995-05-15; SE93/00455
; EARLIER FILING DATE: 1994-11-22
; EARLIER APPLICATION NUMBER: 08/338,501
; EARLIER FILING DATE: 1994-11-22
; EARLIER APPLICATION NUMBER: 08/385,540
; EARLIER FILING DATE: 1995-02-08
; EARLIER APPLICATION NUMBER: 08/486,820
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/600,273
; EARLIER FILING DATE: 1996-02-08

```

```

; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Salmon
;
US-09-220-731-17

Query Match      82.1%; Score 78; DB 4; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.2e-07;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYKXCHSQAHQVLSNS 20
Db 1 IVGGTCKATSAQVLSNS 20

RESULT 12
US-08-978-404B-44
; Sequence 44, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumeri, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: 80801/7090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5968782e
;
US-08-978-404B-44

Query Match      74.7%; Score 71; DB 2; Length 246;
Best Local Similarity 70.0%; Pred. No. 2.9e-05;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYKXCHSQAHQVLSNS 20
Db 24 IVGGTCTPEHSVPVQVLSNS 43

RESULT 13
US-08-956-267A-2

```

```

; Sequence 2, Application US/08956267A
; Patent No. 5945328
; GENERAL INFORMATION:
; APPLICANT: KJELDSEN, Helle Fabricius
; APPLICANT: KJELDSEN, Thomas Borglum
; TITLE OF INVENTION: A Process For Producing Trypsin
; TITLE OF INVENTION: (Trypsinogen)
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: No. 59453280 No. 5945328diak of No. 5945328th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,267A
; FILING DATE: 22-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4500.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-956-467A-2

Query Match 69.5%; Score 66; DB 2; Length 247;
Best Local Similarity 53.6%; Pred. No. 0.00023;
Matches 43; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXHSQAHQVSLNS 20
    ||||| | : |||||
Db 25 IVGGYTCANSPYQVSLNS 44

RESULT 14
US-08-467-155A-10
; Sequence 10, Application US/08467155A
; Patent No. 5736377
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,155A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-628-198-10

Query Match 69.5%; Score 66; DB 2; Length 271;
Best Local Similarity 53.6%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 4; Indels 8; Gaps 1;

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-155A-10

Query Match 69.5%; Score 66; DB 1; Length 271;
Best Local Similarity 53.6%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 4; Indels 8; Gaps 1;

Qy 1 IVGGYX-----CXHXSQAHQVSLNS 20
    ||||| | : |||||
Db 10 IVGGYSSTRYPITCKAYSQPHQVSLNS 37

RESULT 15
US-08-628-198-10
; Sequence 10, Application US/08628198
; Patent No. 5843694
; GENERAL INFORMATION:
; APPLICANT: Band, Vimla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,198
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/467,155
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-628-198-10

Query Match 69.5%; Score 66; DB 2; Length 271;
Best Local Similarity 53.6%; Pred. No. 0.00026;

```

Matches 15; Conservative 1; Mismatches 4; Indels 8; Gaps 1;

OY 1 IVGGYX-----CXHSCAHQVSLNS 20
| | | | |
Db 10 IVGGYSSTRVPIECKAYSPHQVSLNS 37
| | | | |

Search completed: February 12, 2003, 10:30:06
Job time : 5.29851 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:16:06, / Search time 14.1493 Seconds
(without alignment)
291.248 Million cell updates/sec

Title: US-10-036-371-1

Perfect score: 95

Sequence: 1 IVGGYCXKHSHQAHQVSLNS 20

Scoring table: BLASTSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_arChaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_unclassified.*
- 14: sp_vertebrate.*
- 15: sp_viridae.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query #	Score	Match	Length	DB ID	Description
1	80	94.7	219	13	Q91036	Q91036 gadus morhu
2	80	93.7	117	13	Q9PJC6	Q9dfj6 gilllichthys
3	80	93.7	117	13	Q9PJC6	Q93266 pseudopleur
4	79	83.2	240	13	Q9PJC6	Q98tho engraulis j
5	78	82.1	238	13	Q9PJC6	Q9w7g6 parailichth
6	78	82.1	232	13	Q9PJC6	Q9w7g6 parailichth
7	77	81.1	241	13	Q9PJC6	Q9w7g6 engraulis j
8	72	75.8	244	13	Q9PJC6	Q91515 parailichth
9	72	75.8	245	13	Q9PJC6	Q91515 petromyzon
10	71	74.7	247	13	Q9PJC6	Q9w7g6 parailichth
11	69	72.6	217	13	Q9PJC6	Q91515 fuigu rubrip
12	69	72.6	242	13	Q9PJC6	Q92099 paranorche
13	67	70.5	247	11	Q9PJC6	Q9cpn7 mus musculu
14	66	69.5	244	13	Q9PJC6	Q9gpn3 anguilla ja
15	65	68.4	246	11	Q9PJC6	Q92199 mus musculu
16	65	68.4	246	11	Q9PJC6	Q9r017 mus musculu

Result No.	Query #	Score	Match	Length	DB ID	Description
17	64	67.4	246	11	Q9QK99	Q9quk9 mus musculu
18	64	66.3	247	13	Q42158	Q42158 petromyzon
19	63	65.3	247	13	Q42608	Q42608 petromyzon
20	62	65.3	247	13	Q9PJC6	Q9cpn7 mus musculu
21	61	64.2	249	13	Q9PJC6	Q9w6k0 nototenila
22	60	63.2	249	13	Q9PJC6	Q9w6k0 nototenila
23	60	63.2	249	13	Q9PJC6	Q9w6k0 nototenila
24	58	61.1	247	11	Q9D777	Q9w6k0 nototenila
25	57	60.0	250	13	Q91265	Q91265 pseudopleur
26	57	60.0	344	13	Q9W6J9	Q9w6j9 mus musculu
27	52	54.7	48	4	Q9UE84	Q9ue84 dissostichu
28	52	54.7	176	4	Q15098	Q15098 homo sapien
29	52	54.7	180	4	Q96A30	Q96a30 homo sapien
30	52	54.7	195	4	Q07277	Q07277 homo sapien
31	52	54.7	234	4	Q15096	Q15096 homo sapien
32	52	54.7	261	6	Q9N1Q1	Q9n1q1 saguinus oe
33	52	54.7	262	4	Q8T1C8	Q8t1c8 homo sapien
34	52	54.7	235	11	Q63274	Q63274 rattus norv
35	51	53.7	235	11	Q63274	Q63274 rattus norv
36	51	53.7	276	11	Q9QVC2	Q9qyn4 mus musculu
37	50	52.6	17	11	Q9QVC2	Q9qyn4 mus musculu
38	50	52.6	26	11	Q9QVC2	Q9qyn4 mus musculu
39	50	52.6	28	11	Q9QVC2	Q9qyn4 mus musculu
40	48	50.5	24	11	Q9QV67	Q9qyn4 mus musculu
41	46	48.4	239	11	Q63275	Q63275 rattus norv
42	46	48.4	261	11	Q63275	Q63275 rattus norv
43	45	47.4	261	6	Q29474	Q29474 canis fami
44	45	47.4	265	5	Q18488	Q18488 penaeus van
45	45	47.4	270	5	Q8WR11	Q8wr11 paralithode

ALIGNMENTS

Result No.	Query #	Score	Match	Length	DB ID	Description
1	80	94.7	219	13	Q91036	Q91036 gadus morhu
2	80	93.7	117	13	Q9PJC6	Q9dfj6 gilllichthys
3	80	93.7	117	13	Q9PJC6	Q93266 pseudopleur
4	79	83.2	240	13	Q9PJC6	Q98tho engraulis j
5	78	82.1	238	13	Q9PJC6	Q9w7g6 parailichth
6	78	82.1	232	13	Q9PJC6	Q9w7g6 parailichth
7	77	81.1	241	13	Q9PJC6	Q9w7g6 engraulis j
8	72	75.8	244	13	Q9PJC6	Q91515 parailichth
9	72	75.8	245	13	Q9PJC6	Q91515 petromyzon
10	71	74.7	247	13	Q9PJC6	Q9w7g6 parailichth
11	69	72.6	217	13	Q9PJC6	Q91515 fuigu rubrip
12	69	72.6	242	13	Q9PJC6	Q92099 paranorche
13	67	70.5	247	11	Q9PJC6	Q9cpn7 mus musculu
14	66	69.5	244	13	Q9PJC6	Q9gpn3 anguilla ja
15	65	68.4	246	11	Q9PJC6	Q92199 mus musculu
16	65	68.4	246	11	Q9PJC6	Q9r017 mus musculu

Query Match 94.7% Score 90; DB 13; Length 219;
Best Local Similarity 85.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 IVGGYCXKHSHQAHQVSLNS 20

```

Db 20 IVGGYECTPHSQAHQVSLNS 39
||||| |||||||
J. Mar. Biotechnol. 0:0-0(1998).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC
DR EMBL; AF012463; AAC32752.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS0135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; signal.
FT SIGNAL 1 15
FT CHAIN 21 242 TRYPSINOGEN 2.
FT SEQUENCE 242 AA; 26180 MW; 08D2A834FB289080 CRC64;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=21117151; PubMed=11172064;
RX Gracey A.Y., Troll J.V., Somero G.N.;
RT "Hypoxia-induced gene expression profiling in the euryoxic fish
Gillichthys mirabilis."
RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC
DR EMBL; AF266240; AAG13359.1; -.
DR HSSP; P35031; IBIT.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
FT NON_TER 117 117
FT SEQUENCE 117 AA; 13038 MW; 37052EB97322EC56 CRC64;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, PYLORIC CAECA, AND PANCREAS;
RA Douglas S.E., Gallant J.W.;
RT "Isolation of cDNAs for Trypsinogen from the Winter Flounder,
RT Pleuronectes americanus."
Query Match 93.7%; Score 89; DB 13; Length 117;
Best Local Similarity 85.0%; Pred. No. 2:2e-09; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3;
Qy 1 IVGGYCXKXHSQAHQVSLNS 20
||||| |||||||
Db 20 IVGGYECTPHSQAHQVSLNS 39
||||| |||||||
J. Mar. Biotechnol. 0:0-0(1998).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC
DR EMBL; AF012463; AAC32752.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease; signal.
FT SIGNAL 1 15
FT CHAIN 21 242 TRYPSINOGEN 2.
FT SEQUENCE 242 AA; 26180 MW; 08D2A834FB289080 CRC64;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC
DR EMBL; AB041929; BAB40329.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
FT SEQUENCE 240 AA; 26036 MW; 299B119BF071464 CRC64;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
CC
DR EMBL; AB041929; BAB40329.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
FT SEQUENCE 240 AA; 26036 MW; 299B119BF071464 CRC64;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, PYLORIC CAECA, AND PANCREAS;
RA Douglas S.E., Gallant J.W.;
RT "Isolation of cDNAs for Trypsinogen from the Winter Flounder,
RT Pleuronectes americanus."
Query Match 93.2%; Score 79; DB 13; Length 240;
Best Local Similarity 75.0%; Pred. No. 4:1e-07; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 5;
Qy 1 IVGGYCXKXHSQAHQVSLNS 20
||||| |||||||
Db 20 IVGGYECTPHSQAHQVSLNS 39
||||| |||||||
RESULT 5

```

[illegible]

RT "Japanese flounder mRNA for trypsinogen 3."
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.

```

RESULT 11
Q91515 PRELIMINARY; PRT; 237 AA.
ID Q91515 PRELIMINARY;
AC Q91515 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DI 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Trypsinogen (Fragment).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphs; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OK NCBI_Taxid=31033;
CC NCBI_Taxid=31033;
PP SEQUENCE FROM N.A.
RA Wang K., Gan L., Lee I., Roach J., Hood L.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBAJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY

```



```

DR EMBL; U25747; AAA75001.1; -.
DR HSSP; P35031; 1BIT.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; Hydrolyase; Serine protease.
FT NON TER 1
SQ SEQUENCE 237 AA; 25726 MW; 30D2DBAC39080C2 CRC64;

Query Match 72.6%; Score 69; DB 13; Length 237;
Best Local Similarity 70.0%; Pred. No. 3.5e-05;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYCXKXHSQAHQVSLNS 20
Db 16 IVGGYCKRNSVAYQVSLNS 35
||||| : : : : :

RESULT 12
Q82099
Q82099 PRELIMINARY; PRT; 242 AA.
AC Q82099;
DT 01-NOV-1996 (TrEMBLrel. 01; Created)
DT 01-NOV-1996 (TrEMBLrel. 01; Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21; Last annotation update)
DE Trypsin precursor (SC 3.4.21.41)
OS Paratotochenia magellanica (Maori cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Paratotochenia.
OX NCBI_TaxID=37005;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Genicot S., Rencier-Deirue P., Edwards D., Van Beeumen J., Garday C.;
RT "Trypsin and trypsinogen from an antarctic fish: molecular basis of
RT Blochmann adaptation." Acta 1238:45-57(1996).
CC -1- SIMILARITY BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; X82223; CA57701.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; Hydrolyase; Serine protease; Signal.
FT NON TER 1
SQ SEQUENCE 242 AA; 26201 MW; 3F4DF7CEB0C4477C CRC64;

Query Match 72.6%; Score 69; DB 13; Length 242;
Best Local Similarity 70.0%; Pred. No. 3.5e-05;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYCXKXHSQAHQVSLNS 20
Db 21 IVGGKCEPSVSPHQVSLNS 40
||||| : : : : :

RESULT 13
Q8QCN7
Q8QCN7 PRELIMINARY; PRT; 247 AA.
AC Q8QCN7;
DT 01-JUN-2001 (TrEMBLrel. 17; Created)
DT 01-JUN-2001 (TrEMBLrel. 17; Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21; Last annotation update)
DE 1810009J06R1K OR TRYPSINOGEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Kawai J., Shiga A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Arakawa T., Harata A., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bares G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suda K., Toyono K., Wang X.H., Weitz C., Whitaker C., Wilming L.,
RA Yabuchi-Yokoyama A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 403:685-690(2001).
RN 2
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RT segments before gene rearrangement." ;
RL J. Immunol. 166:1771-1780(2001).
CC -1- SIMILARITY BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007406; BAB25018.1; -.
DR HSSP; AK000663; AAB69055.1; -.
DR MEROPS; S01.129; IDPO.
DR WGD; MG1:1920876; 1810009J06R1K.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; Hydrolyase; Serine protease.
SQ SEQUENCE 247 AA; 26503 MW; EDF5F7696833C7BC CRC64;

Query Match 70.5%; Score 67; DB 11; Length 247;
Best Local Similarity 68.4%; Pred. No. 8.8e-05;
Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYCXKXHSQAHQVSLNS 19
Db 24 IVGGYTCPSHVPQVSLNS 42
||||| : : : : :

RESULT 14
Q8QCN3

```

```

ID Q8QW3 PRELIMINARY; PRT; 244 AA.
AC Q8QW3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen (EC 3.4.21.41).
GN TRYP.
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OC Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RN
RP SEQUENCE FROM N.A.
RP TISSUE=PANCREAS;
RA Kurokawa T., Suzuki T., Ohta H., Kagawa H., Tanaka H., Unuma T.;
RT "Expression of pancreatic enzyme genes during the early larval stage
of Japanese eel, Anguilla japonica."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070720; BAB85634.1; -.
KW Hydrolyase.
SQ SEQUENCE 244 AA; 26317 MW; 0EB3B68EB706D52D CRC64;

Query Match 69.5%; Score 66; DB 13; Length 244;
Best Local Similarity 65.0%; Prad.No. 0.00014;
Matches 13; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAHQVSLNS 20
Db ||||| ||||| |||||
21 IVGGYECFPHSQPWQNSLNA 40

RESULT 15
Q9ZIR9 PRELIMINARY; PRT; 246 AA.
ID Q9ZIR9
AC Q9ZIR9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE Trypsinogen (EC 3.4.21.41).
GN TRYGN16 OR TRYPSINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN
RP SEQUENCE FROM N.A.
RP STRAIN=BALB/C;
RA Rowen L., Hood L.;
RT "Comparison between strains Balb/C and 129 in a region of the mouse T
cell receptor beta locus."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN
RP SEQUENCE FROM N.A.
RP STRAIN=BALB/C;
RX MEDLINE=21101195; PubMed=11160223;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RT segments before gene rearrangement."
RL J. Immunol. 166:1771-1780(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF107342; AAC79093.1; -.
DR EMBL; AE006655; AAB69088.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.063; -.
DR MGD; MGI:2148749; Trygn16.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.

```

```

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 246 AA; 26134 MW; 34E173B18CA2F463 CRC64;

Query Match 68.4%; Score 65; DB 11; Length 246;
Best Local Similarity 65.0%; Prad.No. 0.00021;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAHQVSLNS 20
Db ||||| ||||| |||||
24 IVGGYTCRENSVPYQVSLNS 43

Search completed: February 12, 2003, 10:27:17
Job time : 16.1493 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

ON protein - protein search, using sw model

Run on: February 12, 2003, 10:04:45, Search time 2.68657 Seconds
(with alignments)
308.768 Million cell updates/sec

Title: US-10-036-371-1

Perfect score: 95

Sequence: 1 IVG3YCXHXSAHQVSLNS 20

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 40.4

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	90	94.7	241	TRV1_GADMO	P16049 Gadus morhua
2	90	94.7	241	TRVX_GADMO	P16049 Gadus morhua
3	80	84.2	242	TRV1_SALSA	P35031 Gadus salar
4	79	83.2	231	TRV2_SALSA	P35032 Gadus salar
5	71	74.7	246	TRV1_RAT	P00762 rattus norv
6	71	74.7	248	TRV3_CHICK	Q90629 Gallus gall
7	69	72.6	246	TRV4_RAT	P32821 rattus norv
8	69	72.6	246	TRVB_RAT	P32822 rattus norv
9	66	69.5	231	TRVP_PIG	P06871 sus scrofa
10	66	69.5	246	TRV1_CANFA	P06871 canis fami
11	65	68.4	246	TRV2_RAT	P00763 rattus norv
12	65	68.4	247	TRV3_HUMAN	P15951 homo sapien
13	65	68.4	304	TRV4_HUMAN	P35030 homo sapien
14	64	67.4	247	TRV1_HUMAN	P07477 homo sapien
15	63	67.4	247	TRV2_HUMAN	P07478 homo sapien
16	63	66.3	247	TRV2_BOVIN	Q29463 bos taurus
17	63	66.3	247	TRV2_CANFA	Q29462 canis fami
18	62	65.3	247	TRV3_RAT	P08426 rattus norv
19	62	65.3	247	TRV3_CHICK	Q90627 Gallus gall
20	62	65.3	248	TRV2_CHICK	Q90628 Gallus gall
21	62	65.3	248	TRV2_MOUSE	P07146 mus musculu
22	60	63.2	246	TRV3_SALSA	P35033 salmo salar
23	57	60.0	238	TRV3_HUMAN	P17059 xenopus lae
24	56	58.9	244	TRV2_RAT	P20151 homo sapien
25	56	58.9	247	TRV4_RAT	P00764 squalus aca
26	55	57.9	261	TRV2_HUMAN	P35034 pleuronecte
27	54	56.8	229	TRVP_SQUAC	P36373 rattus norv
28	54	56.8	250	TRVP_FLEPL	P36375 rattus norv
29	54	56.8	261	TRV7_RAT	P07288 homo sapien
30	52	54.7	244	KLK4_RAT	P36374 rattus norv
31	52	54.7	261	KLK3_HUMAN	P36375 macaca mula
32	52	54.7	261	KLK3_HUMAN	P36374 rattus norv
33	52	54.7	262	KLK1_HUMAN	P06870 homo sapien

34	51	53.7	259	1	KLK2_RAT	P00759 rattus norv
35	51	53.7	259	1	KLK9_RAT	P07647 rattus norv
36	50	52.6	259	1	KLK6_RAT	P36376 rattus norv
37	50	52.6	261	1	KLK1_RAT	P00758 rattus norv
38	49	51.6	277	1	KLK2_HUMAN	Q90473 homo sapien
39	47	49.5	260	1	KLK8_HUMAN	O60259 homo sapien
40	47	49.5	261	1	KLK1_MOUSE	P15947 mus musculu
41	47	49.5	261	1	KLK2_MOUSE	P04071 mus musculu
42	46	48.4	243	1	TRV1_XENLA	P19799 xenopus lae
43	46	48.4	261	1	KLK8_MOUSE	P07628 mus musculu
44	46	48.4	261	1	KLK6_RAT	P36374 rattus norv
45	46	48.4	261	1	KLK2_MOUSE	P36368 mus musculu

ALIGNMENTS

```
RESULT 1
TRV1_GADMO STANDARD; PRT: 241 AA.
AC P16049; Q91040; Q92156;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin 1 precursor (EC 3.4.21.4).
OS Gadus morhua (Atlantic cod).
OC Rukaryocae; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Actinopterygii; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1] TaxID=8049;
RP SEQUENCE FROM N.A.
RC TISSUE=Pyloric caeca;
RX MEDLINE=94039130; PubMed=8223632;
RA Gudmundsdottir A., Gudmundsdottir E., Oskarsson S., Bjarnason J.B.,
RA Eakin A.E., Craik C.S.;
RT "Isolation and characterization of cDNAs from Atlantic cod encoding
RT two different forms of trypsinogen.";
RL Eur. J. Biochem. 217:1091-1097(1993).
RN [2]
RP SEQUENCE OF 20-58.
RC TISSUE=Pyloric caeca;
RX MEDLINE=8210867; PubMed=2707266;
RA Oskarsson E., Fox J.W., Bjarnason J.B.;
RA "Isolation and characterization of trypsin from the poikilotherm
RA Gadus morhua";
RL Eur. J. Biochem. 180:85-94(1989).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y16566; C345424.1; -
DR FPIR; S01586; S035030.
DR HSPG; P00763; 10P0.
DR MEROPS; S01151; -
DR InterPro; IPR001314; Chymotrypsin
DR Pfam; PF00089; trypsin; 1.
DR InterPro; IPR001254; Ser.protease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02400; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
```

```

FT SIGNAL 1 13
FT PROPEP 14 19
FT CHAIN 20 241
FT ACT_SITE 59 59
FT ACT_SITE 103 103
FT ACT_SITE 195 195
FT ACT_SITE 26 155
FT DISULFID 44 60
FT DISULFID 128 228
FT DISULFID 135 201
FT DISULFID 166 180
FT DISULFID 191 215
FT SITE 189 189
FT CONFLICT 25 25
FT CONFLICT 27 28
FT CONFLICT 43 43
FT CONFLICT 49 52
SQ SEQUENCE 241 AA; 25941 MW; 44EC9A0106AD1A68 CRC64;

Query Match 94.7%; Score 90; DB 1; Length 241;
Best Local Similarity 85.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXKHSHQAHQVSLNS 20
Db 20 IVGGYECTRHSHQAHQVSLNS 39

RESULT 2
TRYX_GADMO
ID TRYX_GADMO STANDARD; PRT; 241 AA.
AC Q91041;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin X precursor (EC 3.4.21.4).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RS SEQUENCE FROM N.A.
RC MEDLINE=96035986; PubMed=8223632;
RA MedlineAccession; PubMedAccession;
RA Ekin A.E., Craik C.S., Gudmundsdottir E., Onkarsson S., Bjarnason J.B.,
RA "Isolation and characterization of cDNAs from Atlantic cod encoding
PT two different forms of trypsinogen";
PL Eur. J. Biochem. 217:1091-1097(1993).
RL 2
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X76887; CA54215.1; -;
CC DR HSP; P00763; LDPO.
CC
CC MEROPS; S01.151; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR001254; Ser.protease_Try.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; Tryp_SPC; 1.
CC DR PROSITE; PS02040; TRYP SIN DOM; 1.
CC DR PROSITE; PS00134; TRYP SIN HIS; 1.
CC DR PROSITE; PS00135; TRYP SIN_SER; 1.

```

```

KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 13
FT PROPEP 14 19
FT CHAIN 20 241
FT ACT_SITE 59 59
FT ACT_SITE 103 103
FT ACT_SITE 195 195
FT ACT_SITE 26 155
FT DISULFID 44 60
FT DISULFID 128 228
FT DISULFID 135 201
FT DISULFID 166 180
FT DISULFID 191 215
FT SITE 189 189
SQ SEQUENCE 241 AA; 25976 MW; 853D7C26BCAF9DD7 CRC64;

Query Match 94.7%; Score 90; DB 1; Length 241;
Best Local Similarity 85.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXKHSHQAHQVSLNS 20
Db 20 IVGGYECTRHSHQAHQVSLNS 39

RESULT 3
TRY1_SALSA
ID TRY1_SALSA STANDARD; PRT; 242 AA.
AC P35031;
AT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin I precursor (EC 3.4.21.4).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RS SEQUENCE FROM N.A.
RC MEDLINE=96035986; PubMed=7556223;
RA MedlineAccession; PubMedAccession;
RA "Molecular cloning and characterization of anionic and cationic
PT variants of trypsin from Atlantic salmon.";
PL Eur. J. Biochem. 232:677-685(1995).
RL 2
CC X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS).
CC Smalas A.O., Hordvik A.;
CC "Structure determination and refinement of benzamide-inhibited
PT trypsin from the North Atlantic salmon (Salmo salar) at 1.82-A
RT resolution.";
RL Acta Crystallogr. D 49:318-330(1993).
RL [3]
CC X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS).
CC MEDLINE=95148588; PubMed=7846025;
CC Smalas A.O., Heimstad E.S., Hordvik A., Willassen N.P., Male R.;
CC "Cold adaptation of enzymes: structural comparison between salmon and
RT bovine trypsins.";
RL Proteins 20:149-166(1994).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70073; CAA49678.1; -
DR PIR; S1778; S1778.
DR NCBI; S1778; S1778.
DR WPO; P45031; AB1.
DR InterPro; IPR008567; 254; Ser Protease_Try.
DR Pfam; PF00089; trypsin_1
DR SMART; SM0020; Tryp_Spc; 1
DR PROSITE; PS50240; TRYPSIN_DOM; 1
DR PROSITE; PS00134; TRYPSIN_HIS; 1
DR PROSITE; PS00135; TRYPSIN_SER; 1
DR Hydrolase; Serine protease; Digestion;
DR Multigenic family.
DR NON_TER 1 1
DR SIGNAL <1 4 POTENTIAL.
DR PROPEP 5 9 ACTIVATION PEPTIDE.
DR CHAIN 10 231 TRYPSIN II.
DR ACT_SITE 93 49 CHARGE RELAY SYSTEM (BY SIMILARITY).
DR ACT_SITE 99 49 CHARGE RELAY SYSTEM (BY SIMILARITY).
DR ACT_SITE 185 185 CHARGE RELAY SYSTEM (BY SIMILARITY).
DR DISULFID 16 145 BY SIMILARITY.
DR DISULFID 34 50 BY SIMILARITY.
DR DISULFID 118 208 BY SIMILARITY.
DR DISULFID 125 181 BY SIMILARITY.
DR DISULFID 152 170 BY SIMILARITY.
DR DISULFID 181 205 BY SIMILARITY.
DR SITE 179 179 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
DR SEQUENCE 231 AA; 24823 MW; C54R1ACFE74FAE18 CRC64;
Query Match 83.2%; Score 79; DB 1; Length 231;
Best local Similarity 75.0%; Pred. No. 1.7e-07;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 IVGGYCXKHSAQHQVSLNS 20
||||| : |||||
DB 10 IVGGYEKAYSQHQVSLNS 29
||||| : |||||

```

ID		STANDARD;	PRT,	246 AA.
OS	PT0762;			
AC				
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	21-JUN-2002 (Rel. 41, Last annotation update)			
DE	Trypsin I, anionic precursor (EC 3.4.21.4) (Pretrypsinogen I).			
DE	TRYI.			
GN	Rattus norvegicus (Rat).			
OR	Kukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Eumetazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
NCBI_TaxId=10116;				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RC	MEDLINE=82265624; PubMed=6996710;			
RA	McDonald R.J., Stray S.J., Swift G.H.;			
RT	"Two similar but nonallelic rat pancreatic trypsins. Nucleotide			
RL	sequences of the cloned cDNAs.";			
RL	J. Biol. Chem. 257:9724-9732(1982).			
[2]				
RP	SEQUENCE FROM N.A.			
OS	MEDLINE=85054880; PubMed=6094547;			
RA	Craik C.S., Choo O.L., Swift G.H., Quinto C., McDonald R.J.,			
RA	Rutter W.J.;			
RT	"Structure of two related rat pancreatic trypsin genes.";			
RL	J. Biol. Chem. 259:14255-14264(1984).			
[3]				
RP	X-RAY CRYSTALLOGRAPHY [2.6 ANGSTROMS].			
RE	MEDLINE=87292123; PubMed=3112942;			
EX	SZANG-SF, Chang P-Y, Chen Y-H, Crawford R M, Friesen-Walker J			

RA Xuong N.-H., Hamlin R., Rutter W.J., Craik C.S.;
 RT "The three-dimensional structure of Asn102 mutant of trypsin: role of
 RL Asp102 in serine protease catalysis.",
 RL Science 237:905-909(1987).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- MISCELLANEOUS: THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR
 CC FORM OF TRYPSIN PRODUCED BY THE ADULT PANCREAS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; V01273; CAA24580.1; -;
 CC EMBL; J00778; AAA98518.1; -;
 CC PIR; A00948; TRT1.
 CC PDB; 1TRM; 15-JUL-93.
 CC PDB; 2TRM; 16-JUL-88.
 CC PDB; 1B8A; 30-APR-94.
 CC PDB; 1BRB; 31-JUL-94.
 CC PDB; 1BRC; 31-MAY-94.
 CC MEROPS; S01.094; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYPSIN; 1.
 CC PROSITE; PS00240; TRYPsin DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC K0M Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 CC Multigene family; 3D structure.
 CC SIGNAL
 CC 1 23 ACTIVATION PEPTIDE.
 CC 2 25 TRYPSIN 1, ANIONIC.
 CC 3 24 TRYPSIN 1, ANIONIC.
 CC 4 246 TRYPSIN 1, ANIONIC.
 CC 5 63 TRYPSIN 1, ANIONIC.
 CC 6 63 TRYPSIN 1, ANIONIC.
 CC 7 107 CHARGE RELAY SYSTEM.
 CC 8 107 CHARGE RELAY SYSTEM.
 CC 9 200 CHARGE RELAY SYSTEM.
 CC 10 200 CHARGE RELAY SYSTEM.
 CC 11 150 CHARGE RELAY SYSTEM.
 CC 12 64 CHARGE RELAY SYSTEM.
 CC 13 233 CHARGE RELAY SYSTEM.
 CC 14 139 CHARGE RELAY SYSTEM.
 CC 15 171 CHARGE RELAY SYSTEM.
 CC 16 196 CHARGE RELAY SYSTEM.
 CC 17 185 CHARGE RELAY SYSTEM.
 CC 18 220 CHARGE RELAY SYSTEM.
 CC 19 194 CHARGE RELAY SYSTEM.
 CC 20 25 CHARGE RELAY SYSTEM.
 CC 21 25 CHARGE RELAY SYSTEM.
 CC 22 28 CHARGE RELAY SYSTEM.
 CC 23 32 CHARGE RELAY SYSTEM.
 CC 24 33 CHARGE RELAY SYSTEM.
 CC 25 37 CHARGE RELAY SYSTEM.
 CC 26 42 CHARGE RELAY SYSTEM.
 CC 27 54 CHARGE RELAY SYSTEM.
 CC 28 55 CHARGE RELAY SYSTEM.
 CC 29 60 CHARGE RELAY SYSTEM.
 CC 30 64 CHARGE RELAY SYSTEM.
 CC 31 70 CHARGE RELAY SYSTEM.
 CC 32 73 CHARGE RELAY SYSTEM.
 CC 33 77 CHARGE RELAY SYSTEM.
 CC 34 78 CHARGE RELAY SYSTEM.
 CC 35 86 CHARGE RELAY SYSTEM.
 CC 36 95 CHARGE RELAY SYSTEM.
 CC 37 98 CHARGE RELAY SYSTEM.
 CC 38 100 CHARGE RELAY SYSTEM.
 CC 39 101 CHARGE RELAY SYSTEM.
 CC 40 105 CHARGE RELAY SYSTEM.
 CC 41 106 CHARGE RELAY SYSTEM.
 CC 42 107 CHARGE RELAY SYSTEM.
 CC 43 127 CHARGE RELAY SYSTEM.
 CC 44 131 CHARGE RELAY SYSTEM.
 CC 45 136 CHARGE RELAY SYSTEM.
 CC 46 138 CHARGE RELAY SYSTEM.
 CC 47 143 CHARGE RELAY SYSTEM.
 CC 48 157 CHARGE RELAY SYSTEM.

FT STRAND 159 165
 FT TURN 168 174
 FT TURN 176 182
 FT TURN 181 186
 FT STRAND 183 186
 FT TURN 189 190
 FT STRAND 194 194
 FT TURN 197 198
 FT TURN 200 201
 FT STRAND 203 206
 FT TURN 207 208
 FT STRAND 209 216
 FT TURN 223 224
 FT STRAND 227 231
 FT TURN 232 234
 FT HELIX 236 244
 FT TURN 245 245
 SQ SEQUENCE 246 AA; 25959 MW; 6AFA0AD11943FB5 CRC64;
 Query Match 74.7%; Score 71; DB 1; Length 246;
 Best Local Similarity 70.0%; Pred. No. 5.6e-06;
 Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 IVGGYKCKXHSQHGVSLNS 20
 Db 24 IVGGITCFHSVPQVSLNS 43
 RESULT 6
 TRYPsin CHICK STANDARD; PRT; 248 AA.
 AC Q90639;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin II-P29 precursor (EC 3.4.21.4).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 CC NCBI_TaxID=9031;
 CC SEQUENCE FROM N.A.
 CC TISSUE=Pancreas;
 CC MEDLINE=95251611; PubMed=7733885;
 CC Wang K., Gan L., Lee I., Hood L.E.;
 CC "Isolation and characterization of the chicken trypsinogen gene
 CC family";
 CC Biochem. J. 307:471-479(1995).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE
 CC LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL; U15157; AAA79914.1; -;
 CC MEROPS; S01.151; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYPSIN; 1.
 CC PROSITE; PS00240; TRYPsin DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.

DR PROSITE, PS00135; TRYPSIN SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 MW Multigene family.

FT SIGNAL 1 16 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 26 248 TRYPSIN II-P29.
 FT PROPEP 17 25
 FT ACT_SITE 26 65 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 109 109 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 42 42 BY SIMILARITY.
 FT DISULFID 134 235 BY SIMILARITY.
 FT DISULFID 141 208 BY SIMILARITY.
 FT DISULFID 173 187 BY SIMILARITY.
 FT DISULFID 198 222 BY SIMILARITY.
 FT SITE 196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 248 AA; 26622 MW; E5E16B07622B588E CRC64;

Query Match 74.7%; Score 71; DB 1; Length 248;
 Best Local Similarity 70.0%; Pred. No. 5.6e-06;
 Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYKXKXHSQAHQVSLNS 20
 DB 26 IVGGTCQEHSPVQVSLNS 45

RESULT 7
 TRVY RAT
 ID TRVY RAT STANDARD; PRT; 246 AA.
 AC P32821;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin V-A precursor (EC 3.4.21.4).
 DE Rattus norvegicus [Rat].
 OS Rattus norvegicus [Rat].
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC SEQUENCE=Pancreas;
 RX MEDLINE=92165057; PubMed=1537555;
 RA Kang J., Wiegand U., Mueller-Hill B.;
 RT "Identification of cDNAs encoding two novel rat pancreatic serine
 RT proteases".
 RL Gene 110:181-187(1992).
 CC -/- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -/- SUBCELLULAR LOCATION: Extracellular.
 CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X59012; G3A41751.1; .
 DR PIR; JQ1471; JQ1471.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.092; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 MW Multigene family.

FT SIGNAL 1 15 ACTIVATION PEPTIDE.
 FT CHAIN 25 246 TRYPSIN V-A.
 FT PROPEP 16 24 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 64 64 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 31 160 BY SIMILARITY.
 FT DISULFID 49 65 BY SIMILARITY.
 FT DISULFID 133 233 BY SIMILARITY.
 FT DISULFID 140 206 BY SIMILARITY.
 FT DISULFID 171 185 BY SIMILARITY.
 FT DISULFID 196 220 BY SIMILARITY.
 FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 246 AA; 26900 MW; 1EB59DB8BA1715 CRC64;

Query Match 72.6%; Score 69; DB 1; Length 246;
 Best Local Similarity 65.0%; Pred. No. 1.3e-05;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYKXKXHSQAHQVSLNS 20
 DB 25 IVGGTCQEHSPVQVSLNA 44

RESULT 8
 TRVY RAT
 ID TRVY RAT STANDARD; PRT; 246 AA.
 AC P32822;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin V-B precursor (EC 3.4.21.4).
 DE Rattus norvegicus [Rat].
 OS Rattus norvegicus [Rat].
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=92165057; PubMed=1537555;
 RA Kang J., Wiegand U., Mueller-Hill B.;
 RT "Identification of cDNAs encoding two novel rat pancreatic serine
 RT proteases".
 RL Gene 110:181-187(1992).
 CC -/- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -/- SUBCELLULAR LOCATION: Extracellular.
 CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X59013; G3A41752.1; .
 DR PIR; JQ1472; JQ1472.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.093; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 MW Multigene family.
 FT SIGNAL 1 15 ACTIVATION PEPTIDE.
 FT CHAIN 25 246 TRYPSIN V-B.

```

FT ACT SITE 64
FT ACT SITE 108
FT ACT SITE 200
FT ACT SITE 200
FT DISULFID 31
FT DISULFID 49
FT DISULFID 133
FT DISULFID 140
FT DISULFID 171
FT DISULFID 196
FT SITE 194
SQ SEQUENCE 246 AA; 26819 MW; 1EB899CA1BAB0025 CRC64;

Query Match 72.6%; Score 69; DB 1; Length 246;
Best Local Similarity 65.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYXXHXSQAHOVSINS 20
    ||||| ||| ||| |||
Db 25 IVGGYTCQHSVPYQVSINA 44
    ||||| ||| ||| |||

RESULT 9
TRYP_PIG STANDARD; PRT; 231 AA.
AC TP00761;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-10.
RA Charles M., Rovey M., Guidoni A.A., Desmuelle P.;
RA "On trypsinogen and trypsin of pig."
RL Biochim. Biophys. Acta 69:115-125(1965).
RN [2]
RP SEQUENCE OF 9-231.
RA MEDLINE=73258692; PubMed=4738933;
RA "Hermanson A., Ericsson L.H., Neurath H., Walsh K.A.;
RA "Purification of the amino acid sequence of porcine trypsin by
RA "petamator analysis."
RL Biochemistry 12:3146-3153(1973).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA Huang Q., Liu S., Tang Y.;
RA MEDLINE=93187998; PubMed=8445634;
RA "Refined 1.6-A resolution crystal structure of the complex formed
RA "between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the
RA "squash family. Detailed comparison with bovine beta-trypsin and its
RA "complex."
RL J. Mol. Biol. 229:1022-1030(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA Huang Q., Liu S., Tang Y., Zeng F., Qian R.;
RA MEDLINE=92201369; PubMed=1551419;
RA "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray
RA "crystal structure of its complex with porcine beta-trypsin."
RL FEBS Lett. 297:143-146(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RA Huang Q., Wang Z., Li Y., Liu S., Tang Y.;
RA MEDLINE=95035057; PubMed=7947985;
RA "Refined 1.8-A resolution crystal structure of the porcine epsilon-
RA "trypsin."
RL Biochim. Biophys. Acta 1209:77-82(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.
RA MEDLINE=97390427; PubMed=9242660;
RA Scubbs M.T., Morenweiser R., Sturzebecher J., Bauer M., Bode W.,
RA Huber R., Piechottka G.P., Matuschiner G., Sommerhoff C.P., Fritz H.,

```

```

RA Auerwald E.A.;
RA "The three-dimensional structure of recombinant leech-derived
RA "trypsin inhibitor in complex with trypsin. Implications for the
RA "structure of human mast cell trypsinase and its inhibition."
RL J. Biol. Chem. 272:19931-19937(1997).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.
RX MEDLINE=98046095; PubMed=9384562;
RA Di Marco S., Priestle J.P.;
RA "Structure of the complex of leech-derived trypsinase inhibitor (LDTI)
RA "with trypsin and modeling of the LDTI-trypsin system."
RL Structure 5:1465-1474(1997).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR PIR; A00947; TRYPGR.
DR PDB; IMCT; 31-JAN-94.
DR PDB; IAKS; 12-FEB-97.
DR PDB; IETI; 07-FEB-95.
DR PDB; IIFX; 21-JAN-96.
DR PDB; ILDT; 30-MAY-96.
DR PDB; IANU; 30-JUL-98.
DR PDB; IAVN; 18-NOV-98.
DR MEROPS; S01.151.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PR00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW 3D-structure.
FT PROPEP 1 8 ACTIVATION PEPTIDE.
FT CHAIN 9 231 TRYPSIN.
FT ACT_SITE 48 48 CHARGE RELAY SYSTEM.
FT ACT_SITE 92 92 CHARGE RELAY SYSTEM.
FT ACT_SITE 185 185 CHARGE RELAY SYSTEM.
FT DISULFID 15 145
FT DISULFID 33 49
FT DISULFID 117 218
FT DISULFID 124 191
FT DISULFID 156 170
FT DISULFID 181 205
FT SITE 179 179
FT VARIANT 20 20
FT STRAND 10 10
FT STRAND 13 14
FT TURN 17 18
FT TURN 21 22
FT STRAND 23 27
FT STRAND 31 39
FT TURN 40 41
FT STRAND 42 45
FT HELIX 47 49
FT STRAND 55 58
FT STRAND 62 62
FT TURN 63 64
FT STRAND 71 80
FT TURN 82 83
FT TURN 86 88
FT TURN 90 91
FT STRAND 94 98
FT STRAND 112 112
FT TURN 120 121
FT STRAND 123 128
FT STRAND 142 142
FT STRAND 144 150
FT HELIX 153 159
FT TURN 161 163
FT TURN 166 167

```

REQUIRED FOR SPECIFICITY (BY SIMILARITY).
I -> V.


```

FT STRAND 168 171
FT TURN 174 175
FT STRAND 179 179
FT TURN 182 183
FT TURN 185 186
FT STRAND 188 191
FT TURN 192 193
FT STRAND 197 200
FT TURN 202 203
FT STRAND 206 209
FT TURN 210 210
FT STRAND 212 216
FT TURN 217 219
FT HELIX 221 230
SQ SEQUENCE 231 AA; 24409 MW; A0A125CF7FC138C2 CRC64;

Query Match 69.5%; Score 66; DB 1; Length 231;
Best Local Similarity 65.0%; Pred. No. 4.5e-05;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYCKXKXHQAHQVSLNS 20
||||| : : : : :
DB 9 IVGGTCRANSIPYQVSLNS 28
||||| : : : : :

RESULT 10
TRY1 CANFA STANDARD; PRT; 246 AA.
AC P06871;
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-JAN-1998 (Rel. 06, Last sequence update)
DE Trypsin, cationic precursor (EC 3.4.21.4).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID:9615;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96284628; PubMed=3841794;
RA Birkby S.D.; Latorge K.S.; Scheele G.;
RA "Data structure, regulation of trypsinogen mRNA translation: full-length
RA mRNA structure encoding two oppositely charged trypsinogen isoenzymes
RA in the dog pancreas."
RL Mol. Cell. Biol. 8:2659-2676 (1988).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isab-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL: W11590; AB030900.1; -.
DR PIR: B24273; TRCC.
DR HSP: P00761; 1PT.
DR MEROPS: S01.151.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRY1_SPC; 1.
DR PROSITE: P500240; TRYPSIN_DOM; 1.
DR PROSITE: P500134; TRYPSIN_HIS; 1.
DR PROSITE: P500135; TRYPSIN_SER; 1.
KM Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
NM Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23 ACTIVATION PEPTIDE.

```

```

FT CHAIN 24 246
FT ACT_SITE 63 63 TRYPSIN, CATIONIC
FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 30 160 BY SIMILARITY.
FT DISULFID 48 64 BY SIMILARITY.
FT DISULFID 132 233 BY SIMILARITY.
FT DISULFID 139 206 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 196 220 BY SIMILARITY.
FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 246 AA; 26170 MW; E9E5A1DE2391BBB CRC64;

Query Match 69.5%; Score 66; DB 1; Length 246;
Best Local Similarity 65.0%; Pred. No. 4.8e-05;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVGGYCKXKXHQAHQVSLNS 20
||||| : : : : :
DB 24 IVGGYTCRNSVPVQVSLNS 43
||||| : : : : :

RESULT 11
TRY2 RAT STANDARD; PRT; 246 AA.
AC P00763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1988 (Rel. 36, Last sequence update)
DE Trypsin II, anionic precursor (EC 3.4.21.4) (Pretcrpsinogen II).
GN TRY2
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054880; PubMed=6094547;
RA Craik C.S.; Choo Q.L.; Swift G.H.; Quinto C.; McDonald R.J.;
RA Rutter W.J.;
RT "Structure of two related rat pancreatic trypsin genes."
RL J. Biol. Chem. 259:14255-14264 (1984).
RN [2]
RP SEQUENCE OF 9-246 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Pancreas;
RC MEDLINE=86285624; PubMed=6896710;
RC "Structure of rat trypsinogen gene."
RT "Structure of rat trypsinogen gene."
RT "Structure of the cloned cDNA", pancreatic trypsinogens. Nucleotide
RT sequences of the cloned cDNA", pancreatic trypsinogens. Nucleotide
RN J. Biol. Chem. 257:9724-9732 (1982).
RN [3]
RX X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
RN MEDLINE=91351998; PubMed=1881877;
RA Earnest T.; Fauman E.; Craik C.S.; Stroud R.;
RT "1.59-A structure of trypsin at 120 K: comparison of low temperature
RT and room temperature structures."
RN Proteins 10:171-187 (1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=96214506; PubMed=8634241;
RX Brinen L.S.; Willett W.S.; Craik C.S.; Flatterick R.J.;
RT "X-ray structures of a designed binding site in trypsin show metal-
RT dependent geometry."
RN J. Mol. Biol. 235:699-7008 (1996).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (see <http://www.iab-sib.ch/annouance/>;
 CC or send an email to license@iab-sib.ch).

CC EMBL; V01274; CAA24581.1; -;
 CC EMBL; L00131; AAA98517.1; -;
 CC EMBL; L00130; AAA98517.1; JOINED.
 CC PIR; A22657; TRT2.
 CC PDB; 1ANB; 01-APR-97.
 CC PDB; 1ANB; 01-APR-97.
 CC PDB; 1ANB; 01-APR-97.
 CC PDB; 1ANB; 01-APR-97.
 CC PDB; 1ANB; 24-DEC-97.
 CC PDB; 1SLU; 11-JUL-96.
 CC PDB; 1SLV; 11-JUL-96.
 CC PDB; 1SLW; 11-JUL-96.
 CC PDB; 1SLX; 11-JUL-96.
 CC PDB; 1DPO; 07-JUL-97.
 CC PDB; 3TGI; 23-DEC-98.
 CC PDB; 3TGU; 23-DEC-98.

CC MEROPS; S01.258; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin_1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYPSIN_DOM; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC HydroLase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 CC Multigene family; 3D-structure.

CC SIGNAL 1 15
 CC PROPEP 16 23 ACTIVATION PEPTIDE.
 CC TRYPSIN_I 24 23 TRYPSIN_I.
 CC ACT_SITE 63 63 CHARGE RELAY SYSTEM.
 CC ACT_SITE 107 107 CHARGE RELAY SYSTEM.
 CC ACT_SITE 200 200 CHARGE RELAY SYSTEM.
 CC DISULFID 30 160 CHARGE RELAY SYSTEM.
 CC DISULFID 48 64
 CC DISULFID 132 233
 CC DISULFID 139 206
 CC DISULFID 171 185
 CC DISULFID 196 220
 CC SITE 194 194
 CC CONFLICT 84 84 N -> D (IN RSP. 1).
 CC CONFLICT 88 88 V -> I (IN RSP. 1).
 CC SEQUENCE 246 AA; 26228 MW; ABD3630809AEE606 CRC64;
 CC REQUIRED FOR SPECIFICITY (BY SIMILARITY).

Query Match 68.4%; Score 65; DB 1; Length 246;
 Best Local Similarity 65.0%; Pred. No. 7.4e-05;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 IVGGYCKXKHSQAHQVSLNS 20
 DB 24 IVGGYTCQENSVPVQVSLNS 43

RESULT 12
 ID TRY3 HUMAN STANDARD; PRT; 247 AA.
 AC P15951;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin III precursor (EC 3.4.21.4).
 GN PRS53 OR TRY3.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=90221895; PubMed=2326201;

RA Tani T., Kawashima I., Mita K., Tokiguchi Y.,
 RT "Nucleotide sequence of the human pancreatic trypsinogen III cDNA."
 RL Nucleic Acids Res. 18:1631-1631(1990).
 CC 1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC 1- SUBCELLULAR LOCATION: Extracellular.
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.iab-sib.ch/annouance/>
 CC or send an email to license@iab-sib.ch).

CC EMBL; X15505; CAA33527.1; -;
 CC PIR; S12764; S12764.
 CC HSSP; P00761; 1EFT.

CC MEROPS; S01.174; -;
 CC Genew; HGNC:9486; PRS53.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin_1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYPSIN_DOM; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
 CC HydroLase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 CC Multigene family.

CC SIGNAL 1 15 ACTIVATION PEPTIDE.
 CC PROPEP 16 23
 CC TRYPSIN_I 24 24 TRYPSIN_I.
 CC ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC DISULFID 30 160 BY SIMILARITY.
 CC DISULFID 48 64 BY SIMILARITY.
 CC DISULFID 139 206 BY SIMILARITY.
 CC DISULFID 171 185 BY SIMILARITY.
 CC SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 CC SEQUENCE 247 AA; 26776 MW; 697DE163F1CEB0D6 CRC64;

Query Match 68.4%; Score 65; DB 1; Length 247;
 Best Local Similarity 65.0%; Pred. No. 7.4e-05;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 IVGGYCKXKHSQAHQVSLNS 20
 DB 24 IVGGYTCQENSVPVQVSLNS 43

RESULT 13
 ID TRY4 HUMAN STANDARD; PRT; 304 AA.
 AC P35030; Q15665; Q9UQV3;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin IV precursor (EC 3.4.21.4) (Brain trypsinogen)
 GN PRS54.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
 RC TISSUE=Brain;
 RX MEDLINE=94123994; PubMed=8294000;
 RA Wiegand U., Corbach S., Minn A., Kang J., Mueller-Hill B.,
 RT "Cloning of the cDNA encoding human brain trypsinogen and
 RT characterization of its product."

RL Gene 136:167-175(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM C).
 RA Fukuko S.;
 CC Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- AUCTIONING PRODUCTS: 3 isoforms; A (shown here), B and C; seem
 CC to be alternative products of alternative splicing.
 CC -!- TISSUE SPECIFICITY: PANCREAS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 11.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC -----
 CC EMBL; X72781; CAB56178.1; -;
 CC EMBL; X71345; CAB50484.1; -;
 CC EMBL; Q45417; BAA08257.1; -;
 CC EMBL; S33496; S33496.
 CC EMBL; F05031; F05031.
 CC EMBL; S331; S331.
 CC
 CC InterPro: IP001314; Chymotrypsin.
 CC InterPro: IP001254; Ser. Protease-Try.
 CC Pfam: PF00089; trypsin.1
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM0020; TRYPSIN.
 CC PROSITE: PS00240; TRYPSIN_DOM; 1.
 CC PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC PROSITE: PS00135; TRYPSIN_SER; FALSE NEG.
 CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 CC Multigene family; Alternative splicing.
 CC SIGNAL 1 ? POTENTIAL.
 CC FT PROPEP ? 80 ACTIVATION PEPTIDE.
 CC FT CHAIN 81 304 TRYPSIN IV.
 CC FT ACT_SITE 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 164 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 264 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT DISULFID 87 217 BY SIMILARITY.
 CC FT DISULFID 105 121 BY SIMILARITY.
 CC FT DISULFID 196 263 BY SIMILARITY.
 CC FT DISULFID 228 242 BY SIMILARITY.
 CC FT DISULFID 253 277 BY SIMILARITY.
 CC FT SITE 251 251 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 CC FT VARSPIC 1 45 MCGDPCPDMWPGGKAVKCGKGLAAARFGRVGGQGRG
 CC FT VARSPIC 1 70 GAGL -> M (IN ISOFORM B).
 CC FT VARSPIC 1 70 MCGDPCPDMWPGGKAVKCGKGLAAARFGRVGGQGRG
 CC FT VARSPIC 1 70 GAGLEHPLIGSTWRRAARDADGCEALGT -> WNPFLILA
 CC FT VARSPIC 1 70 FYGAA (IN ISOFORM C).
 CC FT CONFLICT 89 MISSING (IN REF. 1; CAA50484).
 CC FT SEQUENCE 304 AA; 32499 MW; 4C4316C3F1D0FFC CRC64;
 CC
 CC Query Match 66.4%; Score 65; DB 1; Length 304;
 CC Seq. Loca. Similarity 65.0%; Pred. No. 9.3e-05;
 CC Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC OY 1 IGVGYXXKHSAHOVLNS 20
 CC |||||
 CC DB 81 IGVGYTCENSLPQVSLNS 100
 CC
 CC RESULT 14
 CC TRYL HUMAN
 CC ID TRYL HUMAN STANDARD; PRT; 247 AA.
 CC AC P07477; Q92955; O9HAN6; Q9HAN6; Q9HAN7;
 CC DT 01-APR-1988 (Rel. 07, Created)
 CC DT 01-APR-1988 (Rel. 07, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Trypsin I precursor (EC 3.4.21.4) (Cationic trypsinogen).

GN PSS1 OR TRY1 OR TPPI OR TRYPL.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID:9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:86221712; PubMed:3011602;
 RA Eni M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T.,
 RA Matsubara K.;
 RT "Cloning, characterization and nucleotide sequences of two cDNAs
 RT encoding human pancreatic trypsinogens";
 RL Gene 41:305-310(1986).
 RN [3]
 RP SEQUENCE OF 16-43.
 RX MEDLINE:90091010; PubMed:2598466;
 RA Kimland M., Russick C., Wacker H., Borgstrom A.;
 RT "Immunoreactive anionic and cationic trypsin in human serum";
 RL Clin. Chim. Acta 184:31-46(1989).
 RN [3]
 RP SEQUENCE OF 68-151 FROM N.A., AND VARIANT HPC HIS-122.
 RX MEDLINE:96438847; PubMed:8841182;
 RA Whitcomb D.C., Gorry M.C., Preston R.A., Furey W., Sossenheimer M.J.,
 RA Ulrich C.D., Martin S.P., Gates L.K. Jr., Amann S.T., Toskes P.P.,
 RA Liddle R., McGrath K., Uomo G., Post J.C., Ehrlich G.D.;
 RT "Hereditary pancreatitis is caused by a mutation in the cationic
 RT trypsinogen gene";
 RL Nat. Genet. 14:141-145(1996).
 RN [4]
 RP SEQUENCE OF 15-67 FROM N.A., AND VARIANT HPC GLY-22.
 RX MEDLINE:20399821; PubMed:1093081;
 RA "Chronic pancreatitis, Hoffmeister A., Manns M., Mossner J., Keim V.;
 RA that facilitates trypsin activation";
 RL Gastroenterology 119:461-465(2000).
 RN [5]
 RP SEQUENCE OF 68-151 FROM N.A., AND VARIANTS HPC P-104; C-116 AND F-139.
 RA Teich N., Bauer N., Mossner J., Keim V.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND PHOSPHORYLATION.
 RX MEDLINE:96266496; PubMed:8683601;
 RA Gaboriaud C., Serre L., Guy-Crotte O., Forest E.,
 RA Fontecilla-Camps J.-C.;
 RT "Crystal structure of human trypsin 1: unexpected phosphorylation of
 RT Tyr151";
 RL Mol. Biol. 259:995-1010(1996).
 RN [7]
 RP VARIANTS HPC ILE-29 AND HIS-122.
 RX MEDLINE:9743797; PubMed:922458;
 RA Gorry M.C., Gbazi-Zadeh D., Furey W., Gates L.K. Jr., Preston R.A.,
 RA Aston C.E., Zhang Y., Ulrich C., Ehrlich G.D., Whitcomb D.C.;
 RT "Mutations in the cationic trypsinogen gene are associated with
 RT recurrent acute and chronic pancreatitis";
 RL Gastroenterology 113:1063-1068(1997).
 RN [8]
 RP VARIANT HPC ILE-29.
 RX MEDLINE:98295575; PubMed:9633818;
 RA Teich N., Mossner J., Keim V.;
 RT "Mutations of the cationic trypsinogen in hereditary pancreatitis";
 RL Hum. Mutat. 12:39-43(1998).
 RN [9]
 RP VARIANTS HPC VAL-16 AND HIS-122.
 RX MEDLINE:97315544; PubMed:10361903;
 RA Witt H., Bucher M., Becker M.;
 RT "A signal peptide cleavage site mutation in the cationic trypsinogen
 RT gene is strongly associated with chronic pancreatitis";
 RL Gastroenterology 117:7-10(1999).
 RN [10]
 RP VARIANT HPC ARG-23.
 RX MEDLINE:99219545; PubMed:10204851;
 RA Ferec C., Raguenes O., Salomon R., Roche C., Bernard J.P., Guillot M.,
 RA Quere I., Faure C., Mercier B., Audrezet M.P., Guillausseau P.J.;

5 PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 5.01493 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-1

Perfect score: 95

Sequence: 1 IVGGYCXKHSAQHVSLNS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%

Maximum March 100%

Listing first 45 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	94.7	37	2 S03570	trypsin EC 3.4.21
2	90	94.7	240	2 S39047	trypsin EC 3.4.21
3	90	94.7	241	2 S39048	trypsin EC 3.4.21
4	84	88.4	242	2 S31776	trypsin EC 3.4.21
5	80	84.2	242	2 S31775	trypsin EC 3.4.21
6	79	83.2	231	2 S31778	trypsin EC 3.4.21
7	71	74.7	246	1 TRRT1	trypsin EC 3.4.21
8	71	74.7	248	2 S55066	trypsin EC 3.4.21
9	69	72.6	242	2 S49499	trypsin EC 3.4.21
10	69	72.6	246	2 JQ1472	trypsin EC 3.4.21
11	66	69.5	246	2 JQ1471	trypsin EC 3.4.21
12	66	69.5	231	1 TRGTR	trypsin EC 3.4.21
13	66	69.5	246	1 TRGTR	trypsin EC 3.4.21
14	65	68.4	246	1 TRGTR	trypsin EC 3.4.21
15	65	68.4	247	2 S12764	trypsin EC 3.4.21
16	65	68.4	304	2 S33496	trypsin EC 3.4.21
17	64	67.4	247	1 A25852	trypsin EC 3.4.21
18	64	67.4	247	1 B25852	trypsin EC 3.4.21
19	63	66.3	229	1 TRBTR	trypsin EC 3.4.21
20	63	66.3	247	2 S13813	trypsin EC 3.4.21
21	62	65.3	247	1 TRDG	trypsin EC 3.4.21
22	62	65.3	247	2 A27547	trypsin EC 3.4.21
23	62	65.3	248	2 S55067	trypsin EC 3.4.21
24	60	63.2	246	2 B25528	trypsin EC 3.4.21
25	58	61.1	300	2 A61333	trypsin EC 3.4.21
26	57	60.0	238	2 S31779	trypsin EC 3.4.21
27	56	58.9	247	2 S05494	trypsin EC 3.4.21
28	55	56.4	259	2 I38363	trypsin EC 3.4.21
29	55	57.9	261	2 A29586	tissue kallikrein

ALIGNMENTS

RESULT 1

S03570

trypsin EC 3.4.21.4) - Atlantic cod (fragment)

C:Species: Gadus morhua (Atlantic cod)

C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 30-Sep-1993

C:Accession: S03570

R:Asgeirsson, B.; Fox, J.M.; Bjarnason, J.B.

A:J. Biochem. 180, 85-94, 1989

A:Title: Purification and characterization of trypsin from the polkilotherm Gadus morhua

A:Reference number: S03570; PMID:89210867; PMID:2707266

A:Molecule type: Protein

A:Residues: 357

C:Superfamily: trypsin

C:Keywords: hydrolase, serine proteinase

Query Match

Best Local Similarity 94.7%; Score 90; DB 2; Length 37;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYCXKHSAQHVSLNS 20

Db 1 IVGGYCEAHSAQHVSLNS 20

RESULT 2

S39047

trypsin EC 3.4.21.4) - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C>Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000

C:Accession: S39047

R:Gudmundsdottir, A.; Gudmundsdottir, E.; Oskarsson, S.; Bjarnason, J.B.; Eakin, A.K.; C

Eur. J. Biochem. 217, 1091-1097, 1993

A:Title: Isolation and characterization of cDNAs from Atlantic cod encoding two different

A:Reference number: S39047; PMID:94039130; PMID:8223632

A:Accession: S39047

A:Molecule type: mRNA

A:Residues: 1-240 <GUD>

A:Cross-references: EMBL:X76886; NID:9450517; PIDN:CAA54214.1; PTD:g1334752

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F15-233/Domain: trypsin homology <TRY>

F58,102,194/Active site: His, Asp, Ser #status predicted

Query Match

Best Local Similarity 94.7%; Score 90; DB 2; Length 240;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYCXKHSAQHVSLNS 20

Db 19 IVGGYCEAHSAQHVSLNS 38

```

RESULT 3
S31048
C:Species: EC 3.4.21.4) X - Atlantic cod
C:Species: Gadus morhua (Atlantic cod)
C:Date: 20-May-1994 #sequence_revision 03-Aug-1995 #text_change 20-Jun-2000
C:Accession: S39048
R:Gudmundsdottir, A.; Gudmundsdottir, E.; Oskarsson, S.; Bjarnason, J.B.; Eakin, A.K.; C
A:Title: Isolation and characterization of cDNAs from Atlantic cod encoding two differ
A:Reference number: S39047; MUID:94039130; PMID:8223632
A:Accession: S39048
A:Status: preliminary
A:Keywords: trypsin
A:Molecule type: mRNA
A:Residues: 1-241 <MAL>
A:Cross-references: EMBL:X76887; NID:G450519; PIDN:CAA54215.1; PID:g1334753
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:16-20/Domain: activation peptide #status predicted <MAT>
F:21-235/Domain: trypsin homology <TRY>
F:20-234/Domain: trypsin homology <TRY>
F:59,103,195/Active site: His, Asp, Ser #status predicted

Query Match 94.7%; Score 90; DB 2; Length 241;
Best Local Similarity 85.0%; Pred. No. 1.3e-08;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAQHVSLNS 20
Db 20 IVGGYECYKASQAQHVSLNS 39

RESULT 4
S31776
C:Species: EC 3.4.21.4) IA precursor - Atlantic salmon
C:Species: Salmo salar (Atlantic salmon)
C:Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S66659; S31776
R:Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisaen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of try
A:Reference number: S66657; MUID:96035908; PMID:7556223
A:Accession: S66659
A:Molecule type: mRNA
A:Residues: 1-242 <MAL>
A:Cross-references: EMBL:X70071; NID:G64381; PIDN:CAA49676.1; PID:g64382
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-20/Domain: activation peptide #status predicted <MAT>
F:21-242/Domain: trypsin IA #status predicted <MAT>
F:21-235/Domain: trypsin homology <TRY>
F:127-156,45-61,129-229,136-202,167-181,192-216/Dialfide bonds: #status predicted
F:60,104,196/Active site: His, Asp, Ser #status predicted

Query Match 88.4%; Score 84; DB 2; Length 242;
Best Local Similarity 80.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAQHVSLNS 20
Db 21 IVGGYECYKASQAQHVSLNS 40

RESULT 5
S31775
C:Species: EC 3.4.21.4) I precursor - Atlantic salmon
C:Species: Salmo salar (Atlantic salmon)
C:Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 15-Oct-1999
C:Accession: S66660; S66661; S31775; S31777
R:Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisaen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of try
A:Reference number: S66657; MUID:96035908; PMID:7556223
A:Accession: S66660

```

```

A:Molecule type: mRNA
A:Residues: 1-242 <MAL>
A:Cross-references: EMBL:X70075; NID:G64379; PIDN:CAA49680.1; PID:g64380
A:Experimental source: pancreas
A:Accession: S66661
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 28-32,'A',34-242 <MA2>
A:Cross-references: EMBL:X70072; NID:G64383; PIDN:CAA49677.1; PID:g64384
A:Experimental source: pancreas
A:Note: trypsin IB, probably an allelic variant
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-20/Domain: activation peptide #status predicted <MAT>
F:21-242/Domain: trypsin I #status predicted <MAT>
F:21-235/Domain: trypsin homology <TRY>
F:27-156,45-61,129-229,136-202,167-181,192-216/Dialfide bonds: #status predicted
F:60,104,196/Active site: His, Asp, Ser #status predicted

Query Match 84.2%; Score 80; DB 2; Length 242;
Best Local Similarity 75.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAQHVSLNS 20
Db 21 IVGGYECYKASQAQHVSLNS 40

RESULT 6
S31778
C:Species: EC 3.4.21.4) II precursor - Atlantic salmon (fragment)
C:Species: Salmo salar (Atlantic salmon)
C:Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S66658; S31778
R:Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisaen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of try
A:Reference number: S66657; MUID:96035908; PMID:7556223
A:Accession: S66658
A:Molecule type: mRNA
A:Residues: 1-231 <MAL>
A:Cross-references: EMBL:X70073; NID:G64385; PIDN:CAA49678.1; PID:g64386
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-4/Domain: signal sequence (fragment)
F:5-9/Domain: activation peptide #status predicted <SIG>
F:10-231/Domain: trypsin II #status predicted <MAT>
F:10-224/Domain: trypsin homology <TRY>
F:15-145,34-50,118-218,125-191,156-170,181-205/Dialfide bonds: #status predicted
F:49,93,185/Active site: His, Asp, Ser #status predicted

Query Match 83.2%; Score 79; DB 2; Length 231;
Best Local Similarity 75.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYXCXXHSQAQHVSLNS 20
Db 10 IVGGYECYKASQAQHVSLNS 29

RESULT 7
TRRT1
N:Alternate names: trypsinogen I
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999
C:Accession: B22657; A00948
R:Crain, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A:Title: Structure of rat trypsinogen I
A:Reference number: A22657; MUID:85054880; PMID:6094547
A:Accession: B22657

```


A:Molecule type: DNA
 A:Residues: 1-246 <CRA>
 A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
 A>Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17
 R:Macdonald, R.J.; Stary, S.J.; Swift, G.H.
 J. Biol. Chem. 257, 9724-9732, 1982
 A>Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
 A:Reference numbers: A00946; MUID:82465624; PMID:6896710
 A:Molecule type: mRNA
 A:Residues: 1-246 <CAC>
 A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
 A:Genetics:
 A:Insertions: 14/1; 67/2; 152/1; 197/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-23/Domain: activation peptide #status predicted <APT>
 F:24-246/Product: trypsin I #status predicted <ENZ>
 F:24-239/Domain: trypsin homology <TRY>
 F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F:63,107,200/Active site: His, Asp, Ser #status predicted
 F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu #status predicted
 Query Match 74.7%; Score 71; DB 1; Length 246;
 Best Local Similarity 70.0%; Pred. No. 3.6e-05;
 Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYKXCHSQHQAQVSLNS 20
 Db 24 IVGGYTCPEHSVPQVSLNS 43

RESULT 8
 S55066
 trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
 N:Alternate names: trypsinogen II
 C:Species: Gallus gallus (chicken)
 C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
 C:Accession: S55066; S72347
 R:Wang, K.; Gan, L.; Lee, I.; Hood, L.
 Biochem. J. 307, 471-479, 1995
 A>Title: Isolation and characterization of the chicken trypsinogen gene family.
 A:Reference numbers: S55066; MUID:95251611; PMID:7733885
 A:Molecule type: mRNA
 A:Residues: 1-248 <HANI>
 A:Cross-references: EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g603907
 A:Experimental source: clone 2-P29
 A:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-25/Domain: activation peptide #status predicted <APT>
 F:26-246/Product: trypsin II #status predicted <MAT>
 F:26-241/Domain: trypsin homology <TRY>
 F:65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 74.7%; Score 71; DB 2; Length 248;
 Best Local Similarity 70.0%; Pred. No. 3.6e-05;
 Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYKXCHSQHQAQVSLNS 20
 Db 26 IVGGYTCPEHSVPQVSLNS 45

RESULT 9
 S49489
 trypsin (EC 3.4.21.4) precursor - Paratotothenia magellanica
 C:Species: Paratotothenia magellanica
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
 C:Accession: S49489
 R:Genicot, S.; Rentier-Delrieu, F.; Edwards, D.; van Beumen, J.; Dodson, G.; Gerday, C.
 submitted to the EMBL Data Library, October 1994
 A>Description: Trypsin and trypsinogen from an antarctic fish: molecular basis of cold ac
 A:Reference number: S49489
 A:Accession: S49489
 A:Molecule type: mRNA
 A:Residues: 1-242 <GEN>
 A:Cross-references: EMBL:X82223; NID:g559507; PIDN:CAA57701.1; PID:g559508
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:Key words: signal sequence #status predicted <SIG>
 F:14-242/Product: trypsin #status predicted <MAT>
 F:21-235/Domain: trypsin homology <TRY>

Query Match 72.6%; Score 69; DB 2; Length 242;
 Best Local Similarity 70.0%; Pred. No. 8e-05;
 Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYKXCHSQHQAQVSLNS 20
 Db 21 IVGGKCSQPSQPIQVSLNS 40

RESULT 10
 JQ1472
 trypsin (EC 3.4.21.4) V precursor, b-form - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 22-Jun-1999
 C:Accession: JQ1471; S23784
 R:Kang, J.; Wiegand, U.; Mueller-Hill, B.
 Gene 110, 181-187, 1992
 A>Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.
 A:Reference number: JQ1471; MUID:92165057; PMID:1537555
 A:Accession: JQ1472
 A:Molecule type: mRNA
 A:Residues: 1-246 <KAN>
 A:Cross-references: EMBL:X59013; NID:g57414; PIDN:CAA1752.1; PID:g57415
 A:Experimental source: pancreas
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-24/Domain: activation peptide #status predicted <ACT>
 F:25-246/Product: trypsin V, b-form #status predicted <MAT>
 F:25-239/Domain: trypsin homology <TRY>
 F:31-160,49-65,133-235,140-206,171-185/Disulfide bonds: #status predicted
 F:64,106,200/Active site: His, Asp, Ser #status predicted
 Query Match 72.6%; Score 69; DB 2; Length 246;
 Best Local Similarity 65.0%; Pred. No. 8.2e-05;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 IVGGYKXCHSQHQAQVSLNS 20
 Db 25 IVGGYTCPEHSVPQVSLNA 44

RESULT 11
 JQ1471
 trypsin (EC 3.4.21.4) V precursor, a-form - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 22-Jun-1999
 C:Accession: JQ1471; S23784
 R:Kang, J.; Wiegand, U.; Mueller-Hill, B.
 Gene 110, 181-187, 1992
 A>Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.
 A:Reference number: JQ1471; MUID:92165057; PMID:1537555
 A:Accession: JQ1471
 A:Molecule type: mRNA
 A:Residues: 1-246 <KAN>

A;Cross-references: EMBL:X59012; NID:G57412; PIDN:CAA1751.1; PID:G57413
 A;Exprefamily: source: pancreas
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; protein digestion; serine proteinase; zymogen
 F;1-15/Domain: signal sequence #status predicted <SIG>
 F;16-24/Domain: activation peptide #status predicted <ACT>
 F;25-246/Product: activation V, a-form #status predicted <MAT>
 F;25-239/Domain: trypsin homology <TRY>
 F;31-160,49-65,133-233,140-206,171-185/Disulfide bonds: #status predicted
 F;64,108,200/Active site: His, Asp, Ser #status predicted
 Query Match 72.6%; Score 69; DB 2; Length 246;
 Best Local Similarity 65.0%; Pred. No. 0.2e-05;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYXXKXHSQAHQVSLNS 20
 Db 25 IVGGYTCQHSVPYQVSLNA 44
 ||||| : : : : :
 ||||| : : : : :
 RESULT 12
 TRTPSR
 N;Contains: trypsinogen
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 24-Apr-1984 #sequence_revision 34-Apr-1984 #text_change 31-Mar-2000
 C;Accession: A90641; A90368; A00947
 R;Charles, M.; Soveri, M.; Guidoni, A.; Desnuelle, P.
 Biochim. Biophys. Acta 69, 115-129, 1963
 A;Title: Su le trypsinogene et la trypsine de porc.
 A;Reference number: A90641
 A;Accession: A90641
 A;Molecule type: protein
 A;Residues: 1-10 <CHA>
 R;Hermodon, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
 Biochemistry 12, 3146-3153, 1973
 A;Title: Determination of the amino acid sequence of porcine trypsin by sequenator analysis
 A;Reference number: A90368; PMID:17325662; PMID:4738933
 A;Accession: A90368
 A;Molecule type: protein
 A;Residues: 1-231 <PER>
 A;Position at position 20, Ile and Val occur alternatively
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zymogen
 F;1-231/Product: trypsinogen #status experimental <ZYM>
 F;1-8/Domain: activation peptide #status experimental <MAT>
 F;9-231/Product: trypsin #status experimental <MAT>
 F;9-224/Domain: trypsin homology <TRY>
 F;15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
 F;18,92,185/Active site: His, Asp, Ser #status predicted
 F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
 Query Match 69.5%; Score 66; DB 1; Length 231;
 Best Local Similarity 65.0%; Pred. No. 0.00027;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYXXKXHSQAHQVSLNS 20
 Db 9 IVGGYTCANSPYQVSLNS 28
 ||||| : : : : :
 ||||| : : : : :
 RESULT 13
 TRDSC
 N;Contains: trypsin (EC 3.4.21.4) precursor, cationic - dog
 N;Alternate names: cationic trypsinogen
 C;Species: Canis lupus familiaris (dog)
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
 C;Accession: B26273
 R;Pinkey, S.D.; LaForge, K.S.; Scheele, G.
 Mol. Cell. Biol. 5, 2669-2676, 1985
 A;Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequence
 A;Reference number: A26273; PMID:86284628; PMID:3841794
 A;Accession: B26273

A;Molecule type: mRNA
 A;Residues: 1-246 <PIN>
 A;Cross-references: GB:M11590; NID:G164096; PIDN:AAA30900.1; PID:G164097
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-15/Domain: signal sequence #status predicted <SIG>
 F;16-23/Domain: activation peptide #status predicted <ACT>
 F;24-246/Product: trypsin, cationic #status predicted <ENZ>
 F;24-239/Domain: trypsin homology <TRY>
 F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F;63,107,200/Active site: His, Asp, Ser #status predicted
 F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
 Query Match 69.5%; Score 66; DB 1; Length 246;
 Best Local Similarity 65.0%; Pred. No. 0.00028;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYXXKXHSQAHQVSLNS 20
 Db 24 IVGGYTCRNSVPYQVSLNS 43
 ||||| : : : : :
 ||||| : : : : :
 RESULT 14
 TRRT2
 N;Contains: trypsin (EC 3.4.21.4) II precursor - rat
 A;Alternate names: trypsinogen II
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 18-Jul-1997
 C;Accession: A26857; A00949
 R;Cralk, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
 J. Biol. Chem. 259, 14255-14264, 1984
 A;Title: Structure of two related rat pancreatic trypsin genes.
 A;Reference number: A26857; PMID:65054880; PMID:6094547
 A;Accession: A26857
 A;Molecule type: DNA
 R;Wasmald, R.J. <CR>
 J. Biol. Chem. 257, 9724-9732, 1982
 A;Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
 A;Reference number: A00948; PMID:82265624; PMID:6896710
 A;Accession: A00949
 A;Molecule type: mRNA
 A;Residues: 9-246 <MAC>
 C;Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I mRNA
 C;Genetics:
 A;Introns: 14/1; 67/2
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-15/Domain: signal sequence #status predicted <SIG>
 F;16-23/Domain: activation peptide #status predicted <ACT>
 F;24-246/Product: trypsin II #status predicted <ENZ>
 F;24-239/Domain: trypsin homology <TRY>
 F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
 F;63,107,200/Active site: His, Asp, Ser #status predicted
 F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
 Query Match 68.4%; Score 65; DB 1; Length 246;
 Best Local Similarity 65.0%; Pred. No. 0.00043;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 IVGGYXXKXHSQAHQVSLNS 20
 Db 24 IVGGYTCQNSVPYQVSLNS 43
 ||||| : : : : :
 ||||| : : : : :
 RESULT 15
 S12764
 N;Contains: trypsin (EC 3.4.21.4) III precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
 C;Accession: S12764
 R;Tani, T.; Kawashima, I.; Mita, K.; Takiguchi, Y.
 Nucleic Acids Res. 18, 1631, 1990

A:Title: Nucleotide sequence of the human pancreatic trypsinogen III cDNA.
A:Reference number: S12764; WUID:90221895; PMID:2326201
A:Accession: S12764
A:Molecule type: mRNA
A:Residues: 1-247 <TAN>
A:Cross-references: EMBL:X15505; NID:g37459; PIDN:CAA33527.1; PID:g37460
C:Genetics:
A:Gene: GDB:PRSS3; TRY3
A:Cross-references: GDB:335297
A:Map position: 7q35-7q35
C:Superfamily: trypsin; trypsin homology
F:Keywords: calcium binding; hydrolase; pancreas; protein digestion; serine proteinase;
F:16-21/Domain: signal sequence; #status predicted <Sig>
F:16-21/Domain: signal sequence; #status predicted <Sig>
F:22-247/Product: trypsin; #status predicted <APM>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 68.4%; Score 65; DB 2; Length 247;
Best Local Similarity 65.0%; Pred. No. 0.00043;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 IVGGYXXHXSQAHOVSILNS 20
Db 24 IVGGYTCENSLPYQVSILNS 43
Search completed: February 12, 2003, 10:28:46
Job time : 5.01493 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 5.01493 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-2

Perfect score: 107

Sequence: 1 IVGGYTCGANTVPQVSLNS 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%

Maximum Watch 100%

Listing first 45 summaries

Database : PIR 73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	229	1 TRBOTR	trypsin (EC 3.4.21)
2	96	89.7	231	1 TRBOTR	trypsin (EC 3.4.21)
3	92	86.0	246	1 TRDQC	trypsin (EC 3.4.21)
4	90	84.1	246	1 TRDQC	trypsin (EC 3.4.21)
5	89	83.2	247	2 S13813	trypsin (EC 3.4.21)
6	87	81.3	247	1 TRDG	trypsin (EC 3.4.21)
7	87	81.3	247	2 S12764	trypsin (EC 3.4.21)
8	87	81.3	304	2 S33496	trypsin (EC 3.4.21)
9	85	79.4	246	1 TRRT1	trypsin (EC 3.4.21)
10	85	79.4	247	1 A25852	trypsin (EC 3.4.21)
11	85	79.4	248	2 S55066	trypsin (EC 3.4.21)
12	84	78.5	30	2 A61333	trypsin (EC 3.4.21)
13	84	78.5	247	2 S25852	trypsin (EC 3.4.21)
14	84	78.5	247	2 S25853	trypsin (EC 3.4.21)
15	82	76.6	246	2 T01473	trypsin (EC 3.4.21)
16	82	76.6	246	2 T01471	trypsin (EC 3.4.21)
17	82	76.6	246	2 T01472	trypsin (EC 3.4.21)
18	78.5	73.4	259	2 S13813	trypsin (EC 3.4.21)
19	78	72.9	248	2 S55067	trypsin (EC 3.4.21)
20	74	69.2	20	2 S50023	trypsin-like prote
21	74	69.2	20	2 S50021	trypsin-like prote
22	74	69.2	247	2 S05494	trypsin (EC 3.4.21)
23	70	65.4	231	2 S31778	trypsin (EC 3.4.21)
24	69	64.5	243	2 A35871	trypsin (EC 3.4.21)
25	65	60.7	37	2 S03570	trypsin (EC 3.4.21)
26	62	57.9	238	2 S31779	trypsin (EC 3.4.21)
27	62	57.9	242	2 S31775	trypsin (EC 3.4.21)
28	62	57.9	242	2 S31776	trypsin (EC 3.4.21)
29	60	56.1	48	2 A61331	trypsin (EC 3.4.21)

30 60 56.1 240 2 S39047 trypsin (EC 3.4.21)

31 60 56.1 241 2 S39048 trypsin (EC 3.4.21)

32 60 56.1 241 2 S39048 tissue kallikrein

33 59 55.1 229 1 TRDQ6 trypsin (EC 3.4.21)

34 58 54.2 125 2 A35545 trypsin (EC 3.4.21)

35 58 54.2 104 2 S15395 trypsin (EC 3.4.21)

36 58 54.2 242 2 S49489 trypsin (EC 3.4.21)

37 58 54.2 244 2 A44284 trypsin (EC 3.4.21)

38 58 54.2 259 2 D23863 tissue kallikrein

39 57 53.3 37 2 S23145 tissue kallikrein

40 57 53.3 259 1 KQRTTN tonin (EC 3.4.21)

41 57 53.3 265 1 KQRTTP tissue kallikrein

42 56 52.3 27 2 C61168 coconase (EC 3.4.)

43 56 52.3 43 2 A61168 coconase (EC 3.4.)

44 56 52.3 250 2 T01779 trypsin (EC 3.4.21)

45 56 52.3 250 2 S31384 trypsin (EC 3.4.21)

ALIGNMENTS

RESULT 1

TRBOTR

trypsin (EC 3.4.21.4) precursor - bovine

N:Contains: trypsinogen

C:Species: Bos primigenius taurus (cattle)

C>Date: 24-Apr-1984 #sequence revision 28-Feb-1986 #text change 18-Jul-1997

C:Accession: A90164; A00946; S08774

R:Mikek, O.; Holeysovsky, V.; Tonasek, V.; Sorm, F.

Biochem. Biophys. Res. Commun. 24, 346-352, 1966

A>Title: Covalent structure of bovine trypsinogen. The position of the remaining amides

A:Reference number: A90164; PMID:67168848; PMID:5967094

A:Accession: A90164

A:Residues: types; protein

R:Hartley, R.S., Q., 59-67, 'Q', 69-150, 'N', 152-176, 'N', 178-229 <MK>

Philos. Trans. R. Soc. Lond. B257, 77-87, 1970

A:Contents: annotation; revisions

Biochemistry 14, 1358-1366, 1975

R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.

A>Title: Amino acid sequence of dogfish trypsin.

A:Reference number: A00950; PMID:75146445; PMID:1092332

A:Contents: annotation; revisions

A:Note: the sequence agrees with that shown

R:Bode, W.; Schwager, P.

J. Mol. Biol. 98, 693-717, 1975

A>Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution

A:Reference number: A2359; PMID:76072097; PMID:512

C:Comment: Trypsinogen; X-ray crystallography, binding sites for calcium, substrate, and

C:Comment: Autocatalytic cleavage; cleavage leads to beta-trypsin by releasing a termi-

s pseudofamily; trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-229/Product: trypsinogen #status experimental <2YW>

F:1-6/Domain: activation peptide #status experimental <AP>

F:1-222/Domain: trypsin homology <TRY>

F:7-131,132-229/Product: alpha-trypsin #status experimental <MPT>

F:6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental

F:13-143,31-47,115-216,122-189,154-168,179-203/Diulfide bonds: #status experimental

F:46,90,183/Active site: His, Asp, Ser #status experimental

F:58,60,63,69/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental

F:131-132/Cleavage site: Lys-Ser (autoclytic) #status experimental

Query Match 100.0% Score 107; DB 1; Length 229;

Best Local Similarity 100.0%; Pred. No. 6.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPQVSLNS 20

DB 7 IVGGYTCGANTVPQVSLNS 26

```

RESULT 2
TRYPTIN
  trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
  N/Ontogen: tryptinogen
  C/Species: Sus scrofa domestica (domestic pig)
  C/Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
  C/Accession: A90641; A90368; A90947
  C/Author: M.; Roversy, M.; Guidoni, A.; Desnuelle, P.
  Biochim. Biophys. Acta 69, 115-129, 1963
  A/Title: Su le trypsinogene et la trypsine de porc.
  A/Reference number: A90641
  A/Accession: A90641
  A/Molecule type: protein
  A/Residues: 1-10 <CRA>
  R.Hermanson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
  Biochemistry 12, 3146-3153, 1973
  A/Title: Determination of the amino acid sequence of porcine trypsin by sequenator analysis
  A/Reference number: A90368; MUID:73258692; PMID:4738933
  A/Accession: A90368
  A/Molecule type: protein
  A/Note: at position 20, Ile and Val occur alternatively
  C/Superfamily: trypsin; trypsin homology
  C/Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zymogen
  F.1-231/Product: trypsinogen #status experimental <ZYM>
  F.1-8/Domain: activation peptide #status experimental <APT>
  F.19-231/Product: trypsin #status experimental <MAT>
  F.19-231/Domain: trypsin homology <TRY>
  F.15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
  F.145,92,185/Active site: His, Asp, Ser #status predicted
  F.60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
  Query Match 89.7%; Score 96; DB 1; Length 231;
  Best Local Similarity 85.0%; Pred. No. 3.4e-07;
  Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 9 IVGGYTCANISIPYQVSLNS 28

RESULT 3
TRYPTIN
  trypsin (EC 3.4.21.4) precursor, cationic - dog
  N/Alternate names: cationic trypsinogen
  C/Species: Canis lupus familiaris (dog)
  C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
  C/Accession: B26273
  C/Author: S. D. LaForge, K.S.; Scheele, G.
  J. Biol. Chem. 266, 2669-2676, 1991
  A/Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequence
  A/Reference number: A26273; MUID:86284628; PMID:3841794
  A/Accession: B26273
  A/Molecule type: mRNA
  A/Residues: 1-246 <PTN>
  A/Cross-references: GB:M11590; NID:G164096; PIDN:AAA30900.1; PID:G164097
  C/Superfamily: trypsin; trypsin homology
  C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
  F.1-15/Domain: signal sequence #status predicted <SIG>
  F.16-23/Domain: activation peptide #status predicted <APT>
  F.24-246/Product: trypsin, cationic #status predicted <ENZ>
  F.24-239/Domain: trypsin homology <TRY>
  F.30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
  F.63,107,200/Active site: His, Asp, Ser #status predicted
  F.75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
  Query Match 86.0%; Score 92; DB 1; Length 246;
  Best Local Similarity 85.0%; Pred. No. 1.4e-06;
  Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCANISIPYQVSLNS 28

```

```

Db 24 IVGGYTCRNSVPYQVSLNS 43

RESULT 4
TRYPTIN
  trypsin (EC 3.4.21.4) II precursor - rat
  N/Alternate names: trypsinogen II
  C/Species: Rattus norvegicus (Norway rat)
  C/Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 18-Jul-1997
  C/Accession: A22657; A00949
  C/Author: C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
  J. Biol. Chem. 259, 14255-14264, 1984
  A/Title: Structure of two related rat pancreatic trypsin genes.
  A/Reference number: A22657; MUID:85054880; PMID:6094547
  A/Accession: A22657
  A/Molecule type: DNA
  A/Residues: 1-246 <CRA>
  J. MacDonald, R.J.; Stary, S.J.; Swift, G.H.
  J. Biol. Chem. 257, 9724-9732, 1982
  A/Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
  A/Reference number: A00948; MUID:82265624; PMID:6896710
  A/Accession: A00949
  A/Molecule type: mRNA
  A/Residues: 9-246 <MAC>
  C/Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I mRNA
  C/Genetics: 14/1; 67/2
  A/Introns: 14/1; 67/2
  C/Superfamily: trypsin; trypsin homology
  C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
  F.1-15/Domain: signal sequence #status predicted <SIG>
  F.16-23/Domain: activation peptide #status predicted <APT>
  F.24-239/Product: trypsin II #status predicted <ENZ>
  F.24-239/Domain: trypsin homology <TRY>
  F.30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
  F.63,107,200/Active site: His, Asp, Ser #status predicted
  F.75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
  Query Match 84.1%; Score 90; DB 1; Length 246;
  Best Local Similarity 85.0%; Pred. No. 2.8e-06;
  Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 24 IVGGYTCRNSVPYQVSLNS 43

RESULT 5
TRYPTIN
  trypsin (EC 3.4.21.4) - bovine
  C/Species: Bos primigenius taurus (cattle)
  C/Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
  C/Accession: S13813
  C/Author: R. Huero, I.; Wicker, C.; Guilleto, P.; Toullec, R.; Puigserver, A.
  Eur. J. Biochem. 193, 767-773, 1990
  A/Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsin
  A/Reference number: S13813; MUID:91065383; PMID:1701147
  A/Accession: S13813
  A/Status: preliminary
  A/Molecule type: mRNA
  A/Residues: 1-247 <HUB>
  A/Cross-references: EMBL:X54703; NID:G829; PIDN:CAA38513.1; PID:G930
  C/Superfamily: trypsin; trypsin homology
  C/Keywords: hydrolase; protein digestion; serine proteinase
  F.24-239/Domain: trypsin homology <TRY>
  F.63,107,200/Active site: His, Asp, Ser #status predicted
  Query Match 83.3%; Score 89; DB 2; Length 247;
  Best Local Similarity 80.0%; Pred. No. 3.9e-06;
  Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 24 IVGGYTCRNSVPYQVSLNA 43

```

```

RESULT 6
TRYPSIN
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: A26273
R:Pinsky, S.D.; LaForge, K.S.; Scheele, G.
Mol. Cell. Biol. 5, 2669-2676, 1985
A:Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequence
A:Reference number: A26273; MUID:86284628; PMID:3841794
A:Accession: A26273
A:Molecule type: mRNA
A:Residues: 1-247 <PIN>
A:Cross-references: GB:M1589; NID:9164094; PIDN:AAA10899.1; PID:g164095
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:16-23/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-247/Product: trypsin; anionic #status predicted <EN2>
F:24-239/Domain: trypsin homology <TRY>
F:10-160/48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 81.3%; Score 87; DB 1; Length 247;
Best Local Similarity 80.0%; Pred. No. 7.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPQVSLNS 20
Db 24 IVGGYTCENSLPQVSLNA 43

RESULT 7
TRYPSIN
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
C:Accession: S12764
R:Tani, T.; Kawashima, I.; Mita, K.; Takiguchi, Y.
Nucleic Acids Res. 18, 1631, 1990
A:Title: Nucleotide sequence of the human pancreatic trypsinogen III cDNA.
A:Reference number: S12764; MUID:90221895; PMID:2326201
A:Accession: S12764
A:Molecule type: mRNA
A:Residues: 1-247 <TAN>
A:Cross-references: EMBL:X15505; NID:g37459; PIDN:CAA33527.1; PID:g37460
C:Genetics:
A:Gene: GB:PSS3; TRY3
A:Cross-references: GB:335297
A:Map position: Tq35-Tq35
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase;
F:16-23/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-247/Product: trypsin III #status predicted <MAT>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 81.3%; Score 87; DB 2; Length 247;
Best Local Similarity 80.0%; Pred. No. 7.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPQVSLNS 20
Db 24 IVGGYTCENSLPQVSLNA 43

RESULT 7
TRYPSIN
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999
C:Accession: B22657; A00948
R:Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A:Title: Structure of two related rat pancreatic trypsin genes.
A:Reference number: A22657; MUID:85054880; PMID:6094547
A:Accession: B22657
A:Molecule type: DNA
A:Residues: 1-246 <CRA>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
A:Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 171
R:MacDonald, R.J.; Stary, W.J.; Swift, G.H.
J. Biol. Chem. 259, 9743-9747, 1984
A:Title: Isolation and characterization of the rat pancreatic trypsinogen gene.
A:Reference number: A00948; MUID:82265624; PMID:6896710
A:Accession: A00948
A:Molecule type: mRNA
A:Residues: 1-246 <MAC>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
C:Genetics:
A:Introns: 14/1; 67/2; 152/1; 197/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:16-23/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-246/Product: trypsin I #status predicted <EN2>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 79.4%; Score 85; DB 1; Length 246;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 8
TRYPSIN
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 15-Aug-1997
C:Accession: S33496
R:Wiegand, U.; Corbach, S.; Minn, A.; Kang, J.; Mueller-Hill, B.
submitted to the EMBL Data Library, March 1993
A:Description: Identification, cloning and characterization of a cDNA encoding a human b
A:Reference number: S33496
A:Accession: S33496
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-304 <WIE>
A:Cross-references: EMBL:X72781
C:Genet: GB:PSS4; TRY4
A:Cross-references: GB:335300
A:Map position: Tq35-Tq35
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:81-296/Domain: trypsin homology <TRY>
F:120,164,257/Active site: His, Asp, Ser #status predicted

Query Match 81.3%; Score 87; DB 2; Length 304;
Best Local Similarity 80.0%; Pred. No. 9.7e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPQVSLNS 20
Db 81 IVGGYTCENSLPQVSLNS 100

RESULT 9
TRYPSIN
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999
C:Accession: B22657; A00948
R:Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A:Title: Structure of two related rat pancreatic trypsin genes.
A:Reference number: A22657; MUID:85054880; PMID:6094547
A:Accession: B22657
A:Molecule type: DNA
A:Residues: 1-246 <CRA>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
A:Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 171
R:MacDonald, R.J.; Stary, W.J.; Swift, G.H.
J. Biol. Chem. 259, 9743-9747, 1984
A:Title: Isolation and characterization of the rat pancreatic trypsinogen gene.
A:Reference number: A00948; MUID:82265624; PMID:6896710
A:Accession: A00948
A:Molecule type: mRNA
A:Residues: 1-246 <MAC>
A:Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
C:Genetics:
A:Introns: 14/1; 67/2; 152/1; 197/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:16-23/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-246/Product: trypsin I #status predicted <EN2>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 79.4%; Score 85; DB 1; Length 246;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 IVGGVTCGANTVPYQVSLNS 20
 ||||| :|||||
 DB 24 IVGGVTCPEHSVPYQVSLNS 43
 ||||| :|||||

RESULT 10

A25852
 trypsin (EC 3.4.21.4) I precursor [validated] - human
 N:Alternate names: trypsin, catenonic, trypsinogen I
 C:Species: Homo sapiens (man)
 C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
 C:Accession: A25852; B61066; A43988
 R:Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
 Gene 41, 305-310, 1986
 A:Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human
 A:Reference number: A91544; MUID:86221712; PMID:3011602
 A:Accession: A25852
 A:Molecule type: mRNA
 A:Residues: 1-247
 A:Cross-references: GB:M22612; NID:9521215; PIDN:AAA61231.1; PID:9521216
 R:Kilmand, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
 Clin. Chim. Acta 184, 31-46, 1989
 A:Title: Immunoreactive anionic and cationic trypsin in human serum.
 A:Reference number: A61066; MUID:90091010; PMID:2598466
 A:Accession: B61066
 A:Molecule type: protein
 A:Residues: 6-43
 R:Koivunen, E.; Huhtala, M.L.; Stenman, U.H.
 J. Biol. Chem. 264, 14095-14099, 1989
 A:Title: Human ovarian tumor associated trypsin. Its purification and characterization
 A:Reference number: A43988; MUID:89340515; PMID:2503510
 A:Accession: A43988
 A:Molecule type: protein
 A:Residues: 16-54
 A:Experimental source: mucinous ovarian tumor cyst fluid
 C:Genetics:
 A:Gene: GDB:PRSS1; TRV1
 A:Cross-references: GDB:119620; OMIM:276000
 A:Map position: 7q35-7q35
 A:Note: The human genome contains at least ten trypsin genes or pseudogenes, at least two
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; phosphotriester; protein digestion; serine proteinase, z
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-246/Product: trypsinogen I #status experimental <SIG>
 F:16-23/Domain: activation peptide #status experimental <AP>
 F:224-239/Domain: trypsin I #status predicted <TR>
 F:224-239/Domain: trypsin homology <TR>
 F:30-160-48-64-139-206-171-185-196-220/disulfide bonds: #status predicted
 F:63,107,200/Active site: His, Asp, Ser #status predicted
 F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 79.4%; Score 85; DB 1; Length 247;
 Best Local Similarity 80.0%; Pred. No. 1.6e-05;
 Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGVTCGANTVPYQVSLNS 20
 ||||| :|||||
 DB 24 IVGGVTCPEHSVPYQVSLNS 43
 ||||| :|||||

RESULT 11

S55066
 trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
 N:Alternate names: trypsinogen II
 C:Species: Gallus gallus (chicken)
 C>Date: 03-Mar-1994 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
 C:Accession: S55066; S12347
 R:Wang, K.; Wood, L.
 Biochem. J. 307, 471-486, 1995
 A:Title: Isolation and characterization of the chicken trypsinogen gene family.
 A:Reference number: S55065; MUID:93251611; PMID:7733885
 A:Accession: S55066
 A:Molecule type: mRNA

A:Residues: 1-248 <MAN>
 A:Cross-references: EMBL:U15157; NID:9603906; PIDN:AAA79914.1; PID:9603907
 A:Experimental source: clone 2-P29
 A:Accession: S72337
 A:Molecule type: DNA
 A:Residues: 1-248 <MAN>
 A:Cross-references: EMBL:U15157; NID:9603906; PIDN:AAA79914.1; PID:9603907

A:Experimental source: clone 2-P29
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-25/Domain: activation peptide #status predicted <AP>
 F:26-248/Product: trypsin II #status predicted <TR>
 F:26-241/Domain: trypsin homology <TR>
 F:65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 79.4%; Score 85; DB 2; Length 248;
 Best Local Similarity 80.0%; Pred. No. 1.6e-05;
 Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVGGVTCGANTVPYQVSLNS 20
 ||||| :|||||
 DB 26 IVGGVTCPEHSVPYQVSLNS 45
 ||||| :|||||

RESULT 12

A61333
 trypsin (EC 3.4.21.4) - edible frog (fragment)
 C:Species: Rana esculenta (edible frog)
 C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
 C:Accession: A61333
 R:Pies, W.; Zwilling, R.; Woodbury, R.G.; Neutath, H.
 FEBS Lett. 109, 45-49, 1980
 A:Title: Amino-terminal amino acid sequences and the evolution of frog (Rana esculenta)
 A:Reference number: A61333; MUID:80113255; PMID:6965480
 A:Accession: A61333
 A:Molecule type: protein
 A:Residues: 1-30 <PI>
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 78.5%; Score 84; DB 2; Length 30;
 Best Local Similarity 75.0%; Pred. No. 3e-06;
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGVTCGANTVPYQVSLNS 20
 ||||| :|||||
 DB 1 IVGGVTCPEHSVPYQVSLNS 20
 ||||| :|||||

RESULT 13

B25852
 trypsin (EC 3.4.21.4) II precursor [validated] - human
 N:Alternate names: trypsin 2; trypsin, anionic; trypsinogen II
 C:Species: Homo sapiens (man)
 C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
 C:Accession: B25852; A61066; B43988
 R:Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
 Gene 41, 305-310, 1986
 A:Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human I
 A:Reference number: A91544; MUID:86221712; PMID:3011602
 A:Accession: B25852
 A:Molecule type: mRNA
 A:Residues: 1-247
 A:Cross-references: GB:M27602; NID:9521217; PIDN:AAA61232.1; PID:9521218
 R:Kilmand, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
 Clin. Chim. Acta 184, 31-46, 1989
 A:Title: Immunoreactive anionic and cationic trypsin in human serum.
 A:Reference number: A61066; MUID:90091010; PMID:2598466
 A:Accession: A61066
 A:Molecule type: protein
 A:Residues: 16-39; X' 41-42, XXXX', 47-49 <KIM>
 R:Koivunen, E.; Huhtala, M.L.; Stenman, U.H.


```

A.Molecule type: mRNA
A.Residues: 1-246 ->STE
A.Cross-references: GB:X04574; PID:954918; PID:CA28243.1; PID:954919
C.Keywords: trypsin; trypsin homology
C.Keywords: calcium binding; hydrolase, protein digestion; serine proteinase
F.24-246/Function: Predicted ->SIG
F.24-246/Product: trypsin #status predicted <MAT>
F.24-239/Domain: trypsin homology <TRY>
F.30-160,148-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F.163,107,200/Active site: His, Asp, Ser #status predicted
F.75,77,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match      76.6%      Score 82;   DB 2;   Length 246;
      Local Similarity 75.0%      Pred. No. 4,6e-05;
Matches 15;      Conservative 3;      Mismatches 2;      Indels 0;      Gaps 0;

QY  1  IVGGYTCGANTVPYQVSLN 20
      |||||
DB   24  IVGGYTCRSSVPYQVSLNA 43
      |||||

Search completed: February 12, 2003, 10:28:47
Job time : 6.01493 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using pw model

Run on: February 12, 2003, 10:04:45 ; Search time 2.68657 Seconds
(without alignments)
308.768 Million cell updates/sec

Title: US-10-036-371-2

Perfect score: 107

Sequence: 1 IVGGYTCGANTVPQVSLNS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	243	1 TRV1_BOVIN	P00760 bos taurus
2	96	89.7	231	1 TRYP_PIG	P00761 sus scrofa
3	92	86.0	246	1 TRV1_CANFA	P06871 canis famill
4	89	84.1	246	1 TRV2_RAT	P00763 rattus norv
5	89	84.1	246	1 TRV2_RAT	P00763 rattus norv
6	87	81.3	247	1 TRV2_CHICK	P00763 rattus norv
7	87	81.3	247	1 TRV2_CHICK	P00763 rattus norv
8	87	81.3	247	1 TRV2_CHICK	P00763 rattus norv
9	85	79.4	246	1 TRV1_HUMAN	P00763 rattus norv
10	85	79.4	247	1 TRV1_HUMAN	P00763 rattus norv
11	85	79.4	248	1 TRV3_CHICK	P00629 gallus gall
12	84	78.5	244	1 TRV2_XENLA	P00059 xenopus lae
13	84	78.5	247	1 TRV2_HUMAN	P00748 homo sapien
14	84	78.5	247	1 TRV3_RAT	P08426 rattus norv
15	82	76.6	246	1 TRV2_MOUSE	P07146 mus musculu
16	82	76.6	246	1 TRV2_MOUSE	P07146 mus musculu
17	82	76.6	246	1 TRV2_MOUSE	P07146 mus musculu
18	82	76.6	246	1 TRV2_MOUSE	P07146 mus musculu
19	78	72.9	248	1 TRV1_CHICK	P00627 gallus gall
20	78	72.9	248	1 TRV2_CHICK	P00628 gallus gall
21	74	69.2	247	1 TRV4_RAT	P12788 rattus norv
22	74	69.2	247	1 TRV2_SALSA	P12788 rattus norv
23	69	65.4	231	1 TRV3_SALSA	P12788 rattus norv
24	62	57.9	238	1 TRV3_SALSA	P12788 rattus norv
25	62	57.9	242	1 TRV1_SALSA	P12788 rattus norv
26	60	56.1	241	1 TRV1_GADMO	P16049 gadus morhu
27	60	56.1	241	1 TRV3_GADMO	P16049 gadus morhu
28	59	55.1	229	1 TRV1_RAT	P06373 rattus norv
29	58	54.2	244	1 TRYP_SOAC	P00764 squallus aca
30	58	54.2	244	1 TRV3_RAT	P06373 rattus norv
31	57	53.3	259	1 TRV2_RAT	P00767 rattus norv
32	57	53.3	259	1 TRV2_RAT	P00767 rattus norv
33	56	52.3	14	1 COCO_LIMPO	P05586 limulus pol

34	56	52.3	250	1 TRYP_PLEPL	P35034 pleuroctene
35	56	52.3	259	1 KLK8_RAT	P36376 rattus norv
36	55	51.4	280	1 TRV2_DROME	P42280 drosophila
37	54	50.5	261	1 KLK1_MOUSE	P15947 mus musculu
38	53	49.5	256	1 TRVE_DROER	P54627 drosophila
39	53	49.5	256	1 TRVE_DROME	P35005 drosophila
40	53	49.5	261	1 KLK8_MOUSE	P07628 mus musculu
41	53	49.5	261	1 KLK8_RAT	P36374 rattus norv
42	53	49.5	261	1 KLK8_MOUSE	P36368 mus musculu
43	53	49.5	261	1 KLK2_MOUSE	P04071 mus musculu
44	53	49.5	261	1 KLK2_MOUSE	P36369 mus musculu
45	52	48.6	259	1 DRF3_DERFA	F49275 dermatophag

ALIGNMENTS

RESULT 1	
ID TRV1_BOVIN	
AC P00760;	
DT 21-JUL-1986 (Rel. 01, Created)	
DT 01-NOV-1997 (Rel. 35, Last sequence update)	
DT 15-JUN-2002 (Rel. 41, Last annotation update)	
DE Trypsinogen, cationic precursor (EC 3.4.21.4) (Beta-trypsin)	
DE (Fragment).	
OS Bos taurus (Bovine).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC Bovidae; Bovinae; Bos.	
OX NCBI_TaxID=9913;	
RN 111	
RC TISSUE=Pancreas;	
RA Okajima T., Maniwa M., Nagao S., Fujikawa H., Goto S.;	
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.	
RN [2]	
RP SEQUENCE OF 15-243, AND DISULFIDE BONDS.	
RX MEDLINE=67168848; PubMed=5967094;	
RA Mike O., Holeyovsky V., Tomasek V., Sorm F.;	
RT "Covalent structure of bovine trypsinogen. The position of the	
RT remaining amides."	
RL Biochem. Biophys. Res. Commun. 24:346-352(1966).	
RN [3]	
RP REVISIONS.	
RX MEDLINE=72035053; PubMed=4399051;	
RA Hattiley B.S.;	
RT "Homologues in serine proteinases."	
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).	
RN [4]	
RP REVISIONS.	
RX MEDLINE=75146445; PubMed=1092332;	
RA Titani K., Ericsson L.H., Neurath H., Walsh K.A.;	
RT "Amino acid sequence of dogfish trypsin."	
RL Biochemistry 14:1358-1366(1975).	
RN [5]	
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.	
RX MEDLINE=76072097; PubMed=512;	
RA Rode W., Schwager P.;	
RT "The refined crystal structure of bovine beta-trypsin at 1.8-A	
RT resolution. II. Crystallographic refinement, calcium binding site,	
RT benzamide binding site and active site at pH 7.0."	
RL J. Mol. Biol. 98:693-717(1975).	
RN [6]	
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).	
RX MEDLINE=77112431; PubMed=556951;	
RA Kosiakoff A.A., Chambers J.L., Kay L.M., Stroud R.M.;	
RT "Structure of bovine trypsinogen at 1.9-A resolution."	
RL Biochemistry 16:654-664(1977).	
RN [7]	
RP DISULFIDE BONDS.	
RX MEDLINE=66079271; PubMed=5892911;	
RA Kauffman D.L.;	

"The disulphide bridges of trypsin.",
 J. Mol. Biol. 12:929-932(1965).
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- SUBCELLULAR LOCATION: Extracellular.
 CC -|- TISSUE SPECIFICITY: SYNTHESIZED IN THE ACINAR CELLS OF THE
 CC PANCREAS.
 CC -|- PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY
 CC RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER
 CC LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190
 CC YIELDS PSEUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- DATABASE: NMS-Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/T/TRY.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D38507; BAA07516.1; -;
 DR PIR; A00946; TRBOTR.
 DR PDB; 2TGA; 09-APR-85.
 DR PDB; 1TGB; 14-MAR-85.
 DR PDB; 1TGC; 09-APR-85.
 DR PDB; 2TGD; 07-MAY-86.
 DR PDB; 1TGN; 22-OCT-84.
 DR PDB; 2TGP; 14-MAR-85.
 DR PDB; 1TGS; 14-MAR-85.
 DR PDB; 1TGT; 15-OCT-90.
 DR PDB; 2TGU; 09-APR-85.
 DR PDB; 2TLD; 15-OCT-92.
 DR PDB; 2TLE; 15-MAR-82.
 DR PDB; 2TFI; 15-JUL-91.
 DR PDB; 2TFJ; 15-MAR-85.
 DR PDB; 4TFI; 08-NOV-85.
 DR PDB; 1TFO; 14-MAR-85.
 DR PDB; 1TTP; 16-APR-87.
 DR PDB; 1TTP; 15-APR-91.
 DR PDB; 3PTB; 14-MAR-85.
 DR PDB; 2PTC; 14-MAR-85.
 DR PDB; 2PTN; 15-APR-91.
 DR PDB; 1GBT; 31-JAN-94.
 DR PDB; 1PTA; 15-JUL-92.
 DR PDB; 1PTB; 07-JUL-97.
 DR PDB; 1PTC; 09-APR-85.
 DR PDB; 1PTD; 15-JUL-92.
 DR PDB; 1PTG; 31-JAN-94.
 DR PDB; 1PPE; 31-JAN-94.
 DR PDB; 1PPH; 31-JAN-94.
 DR PDB; 1SMF; 31-JUL-94.
 DR PDB; 1TNG; 30-NOV-94.
 DR PDB; 1TNI; 30-NOV-94.
 DR PDB; 1TNN; 30-NOV-94.
 DR PDB; 1TNK; 30-NOV-94.
 DR PDB; 1TNL; 30-NOV-94.
 DR PDB; 1BTP; 29-JAN-96.
 DR PDB; 1BTW; 15-OCT-95.
 DR PDB; 1BTX; 15-OCT-95.
 DR PDB; 1BTY; 15-OCT-95.
 DR PDB; 1BTZ; 15-OCT-95.
 DR PDB; 1JRS; 14-OCT-96.
 DR PDB; 1JRT; 14-OCT-96.
 DR PDB; 1MAX; 14-OCT-96.
 DR PDB; 1MAY; 14-OCT-96.
 DR PDB; 1MTS; 20-AUG-97.
 DR PDB; 1MTU; 12-NOV-97.
 DR PDB; 1MTV; 12-NOV-97.
 DR PDB; 1MTW; 12-NOV-97.
 DR PDB; 1TPS; 26-JAN-95.
 DR PDB; 1TVN; 26-JAN-95.
 DR PDB; 1TAM; 24-JUN-97.
 DR PDB; 1A07; 25-FEB-98.
 DR PDB; 1TIO; 23-SEP-98.
 DR PDB; 2TIO; 30-SEP-98.
 DR PDB; 1XUF; 16-DEC-98.
 DR PDB; 1XUG; 16-DEC-98.
 DR PDB; 1XUH; 11-NOV-98.
 DR PDB; 1XUI; 11-NOV-98.
 DR PDB; 1XUJ; 11-NOV-98.
 DR PDB; 1XUK; 11-NOV-98.
 DR PDB; 1XUL; 14-OCT-98.
 DR PDB; 1AUJ; 13-JAN-99.
 DR PDB; 1A28; 13-JAN-99.
 DR PDB; 1BJU; 13-JAN-99.
 DR PDB; 1BJV; 13-JAN-99.
 DR PDB; 2BZA; 23-MAR-99.
 DR PDB; 1CES; 23-MAR-99.
 DR MEROPS; S01.151; -;
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW 3D-structure.
 FT ND-TER 1 1
 FT SIGNAL <1 14
 FT PROPEP 15 20 ACTIVATION PEPTIDE.
 FT CHAIN 21 243 TRYPSIN, CATIONIC.
 FT CHAIN 21 145 ALPHA-TRYPSIN CHAIN 1.
 FT CHAIN 146 243 ALPHA-TRYPSIN CHAIN 2.
 FT CA-BIND 72 82
 FT BINDING 191 192
 FT BINDING 194 195
 FT BINDING 197 197
 FT DISULFID 27 127
 FT DISULFID 27 127
 FT DISULFID 125 221
 FT DISULFID 136 203
 FT DISULFID 168 182
 FT DISULFID 193 217
 FT ACT_SITE 60 60
 FT ACT_SITE 104 104
 FT ACT_SITE 197 197
 FT SITE 191 191
 FT SITE 191 191
 FT STRAND 22 22
 FT STRAND 25 26
 FT TURN 29 30
 FT TURN 33 34
 FT TURN 35 39
 FT STRAND 43 51
 FT STRAND 52 53
 FT STRAND 54 57
 FT STRAND 59 61
 FT STRAND 67 70
 FT STRAND 75 76
 FT STRAND 83 92
 FT TURN 94 95
 FT TURN 97 97
 FT TURN 98 101
 FT STRAND 102 102
 FT TURN 103 103
 FT STRAND 106 110
 FT STRAND 124 124
 FT TURN 132 133
 FT STRAND 135 140
 FT STRAND 156 162
 FT STRAND 165 171
 FT TURN 173 174
 FT TURN 178 179
 FT STRAND 180 183
 FT TURN 186 187

```

Query Match      100.0%; Score 107; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGSYTCGANTPVQVSLNS 20
Db 21 IVGSYTCGANTPVQVSLNS 40

RESULT 2
ID TRYP.PIG STANDARD; PRT; 231 AA.
AC P00761;
DT 21-JUL-1986 (Rel. 01. Created)
DT 21-JUL-1986 (Rel. 01. Last sequence update)
DT 15-JUN-2002 (Rel. 41. Last annotation update)
DE Trypsin precursor (EC 3.4.21.4).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-10.
RA Charles W., Rovey M., Guidoni A.A., Desmuelle P.;
RA On trypsinogen and trypsin of pig.;
RA Biochim. Biophys. Acta 65:115-125(1963).
RN [2]
RP SEQUENCE OF 9-231.
RA MEDLINE=73258692; PubMed=4738913;
RA Hermodson M.A., Ericsson L.H., Neurath H., Walsh K.A.;
RT "Determination of the amino acid sequence of porcine trypsin by
RT sequenator analysis.";
RL Biochemistry 12:3146-3153(1973).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA Huang Q., Liu S., Tang Y.;
RA MEDLINE=93187998; PubMed=8445634;
RT "Refined 1.6-A resolution crystal structure of the complex formed
RT between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the
RT squash family. Detailed comparison with bovine beta-trypsin and its
RT complex.";
RN [4]
RP Mol. Biol. 229:1022-1030(1993).
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA Huang Q., Liu S., Tang Y., Zeng F., Qian P.;
RA MEDLINE=92201369; PubMed=9551419;
RT "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray
RT crystal structure of its complex with porcine beta-trypsin.";
RL FEBS Lett. 297:143-146(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RA Huang Q., Wang Z., Li Y., Liu S., Tang Y.;
RA MEDLINE=95035057; PubMed=7947985;
RT "Refined 1.8-A resolution crystal structure of the porcine epsilon-
RT trypsin.";
RN [6]
RP Biochim. Biophys. Acta 1209:77-82(1994).
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.
RA MEDLINE=9394827; PubMed=8242600;
RA Stubbins M.T., Mawlaei K., Sturzebecher J., Bauer M., Bode W.,
RA Huber R., Piechottka G.F., Matschiner G., Sommerhoff C.P., Fritz H.,
RA Auerwald E.A.;
RT "The three-dimensional structure of recombinant leech-derived
RT trypsinase inhibitor in complex with trypsin. Implications for the
RT structure of human mast cell tryptase and its inhibition.";
RL J. Biol. Chem. 272:19931-19937(1997).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.
RA MEDLINE=98046095; PubMed=9384562;
RA di Marco S., Priestle J.P.;
RT "Structure of the complex of leech-derived trypsinase inhibitor (LDTI)
RT with trypsin and modeling of the LDTI-tryptase system.";

```

```

Structure 5:1465-1474(1997).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR PIR: A00947; TRPGTR.
DR PDB: 1MCT; 31-JAN-94.
DR PDB: 1AKS; 12-FEB-97.
DR PDB: 1EPT; 07-FEB-95.
DR PDB: 1TFX; 21-JAN-98.
DR PDB: 11DT; 20-MAY-98.
DR PDB: 1ANI; 01-JUL-98.
DR PDB: 1AWV; 18-NOV-98.
DR PDB: 1AWV; 18-NOV-98.
DR MEDS: S01; 15-NOV-98.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM0020; TRYPSIN; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW 3D-structure.
FT PROPEP 1 8 ACTIVATION PEPTIDE.
FT CHAIN 9 231 TRYPSIN.
FT ACT_SITE 48 48 CHARGE RELAY SYSTEM.
FT ACT_SITE 92 92 CHARGE RELAY SYSTEM.
FT ACT_SITE 185 185 CHARGE RELAY SYSTEM.
FT DISULFID 33 149.
FT DISULFID 117 218.
FT DISULFID 124 191.
FT DISULFID 156 170.
FT DISULFID 181 205.
FT SITE 179 179.
FT VARIAT 20 20 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
FT STRAND 10 10 I -> V.
FT STRAND 13 14.
FT TURN 17 18.
FT TURN 21 22.
FT STRAND 23 27.
FT STRAND 31 39.
FT TURN 40 41.
FT STRAND 42 45.
FT STRAND 47 48.
FT STRAND 55 58.
FT STRAND 62 62.
FT TURN 63 64.
FT STRAND 71 80.
FT TURN 82 83.
FT TURN 86 88.
FT TURN 90 91.
FT STRAND 94 98.
FT STRAND 112 112.
FT TURN 120 121.
FT STRAND 123 128.
FT STRAND 142 142.
FT STRAND 144 150.
FT STRAND 153 159.
FT TURN 161 163.
FT TURN 166 167.
FT TURN 168 171.
FT TURN 174 174.
FT STRAND 179 179.
FT TURN 182 183.
FT TURN 185 186.
FT STRAND 188 191.
FT TURN 192 193.
FT STRAND 194 202.
FT STRAND 207 207.
FT TURN 208 209.
FT STRAND 210 210.

```

```

FT STRAND 212 216
FT FT HELIX 217 219
FT FT HELIX 221 230
SQ SEQUENCE 231 AA; 24409 MW; A0A125CF7C138C2 CRC64;

Query Match 89.7%; Score 96; DB 1; Length 231;
Best Local Similarity 85.0%; Pred. No. 7e-08;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPYQVSLNS 20
DB 9 IVGGYTCANNSIPYQVSLNS 28

RESULT 3
ID TRY1 CANFA STANDARD; PRT; 246 AA.
AC P06871;
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin, cationic precursor (EC 3.4.21.4).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284628; PubMed=3841794;
RA Pinsky S.D., Laforce K.S., Scheele G.;
RT "Differential regulation of trypsinogen mRNA translation: full-length
RT mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
RT in the dog pancreas.";
RL Mol. Cell. Biol. 5:269-2676(1985).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11590; AAA30900.1; -.
CC DR PIR; B26273; TRDSC.
CC DR HSP; P00761; LEPT.
CC DR MEROPS; S01.151; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR001254; Ser. protease_Try.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; TRY1_SPC; 1.
CC DR PROSITE; PS50240; TRYPsin DOM; 1.
CC DR PROSITE; PS00134; TRYPsin_HIS; 1.
CC DR PROSITE; PS00135; TRYPsin_SER; 1.
CC KW Hydroxylase, Serine protease; digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23 ACTIVATION PEPTIDE.
FT CHAIN 24 246 TRYPSIN, CATIONIC.
FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 30 160 BY SIMILARITY.
FT DISULFID 48 64 BY SIMILARITY.
FT DISULFID 132 233 BY SIMILARITY.
FT DISULFID 139 206 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 196 220 BY SIMILARITY.
FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).

SQ SEQUENCE 246 AA; 26170 MW; E9E5A1DE2391BBB CRC64;

Query Match 86.0%; Score 92; DB 1; Length 246;
Best Local Similarity 85.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPYQVSLNS 20
DB 24 IVGGYTCRSNVPYQVSLNS 43

RESULT 4
ID TRY2 RAT STANDARD; PRT; 246 AA.
AC P00763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II, anionic precursor (EC 3.4.21.4) (Pretypsinogen II).
GN TRY2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054880; PubMed=6094547;
RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
RA Rutter W.J.;
RT "Structure of two related rat pancreatic trypsin genes.";
RL J. Biol. Chem. 259:14255-14264(1984).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC SEQUENCE OF 9-246 FROM N.A.
CC RC STRAIN-Sprague-Dawley; TISSUS=Pancreas;
CC RC MEDLINE=82265624; PubMed=689710;
CC RA McDonald R.J., Scary S.J., Swift G.H.;
CC RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
CC sequences of the cloned cDNAs.";
CC RL J. Biol. Chem. 257:9724-9732(1982).
CC
CC RP X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
CC RX MEDLINE=9135198; PubMed=1881877;
CC RA Earnest T., Pauman E., Craik C.S., Stroud R.;
CC RT "1.59-A structure of trypsin at 120 K: comparison of low temperature
CC and room temperature structures.";
CC RL Proteins 10:171-187(1991).
CC
CC RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
CC RX MEDLINE=96214506; PubMed=8634241;
CC RA Brinen L.S., Willett W.S., Craik C.S., Fletterick R.J.;
CC RT "X-ray structures of a designed binding site in trypsin show metal-
CC dependent geometry.";
CC RL Biochemistry 35:5999-6009(1996).
CC CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC CC -1- SUBCELLULAR LOCATION: Extracellular.
CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01274; CAA24581.1; -.
CC DR EMBL; L00131; AAA98517.1; -.
CC DR EMBL; L00130; AAA98517.1; JOINED.
CC DR PIR; A22657; TRT2.
CC DR PDB; 1ANB; 01-APR-97.
CC DR PDB; 1ANC; 01-APR-97.
CC DR PDB; 1AND; 01-APR-97.
CC DR PDB; 1ANE; 01-APR-97.

```

```

DR PDB: 1AMH; 24-DEC-97.
DR PDB: 1SLU; 11-JUL-96.
DR PDB: 1SLV; 11-JUL-96.
DR PDB: 1SLW; 11-JUL-96.
DR PDB: 1SLX; 11-JUL-96.
DR PDB: 1DPO; 07-JUL-97.
DR PDB: 1TGI; 23-DEC-98.
DR PDB: 1TGF; 23-DEC-98.
DR MEROPS: S01.258-98.
DR MEROPS: S01.258-98.
DR InterPro: IPR001114; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: P000722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
Multigene family; 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT SITE 84 84
FT CONFLICT 88 88
FT CONFLICT 88 88
SQ SEQUENCE 246 AA; 26228 MW; A8D3630809AE8606 CRC64;

Query Match 84.1%; Score 90; DB 1; Length 246;
Best Local Similarity 85.0%; Pred. No. 6.5e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPYQVSLNS 20
DB 24 IVGGYTCGANTVPYQVSLNS 43

RESULT 5
ID TRY2 BOVIN STANDARD; PRT; 247 AA.
AC Q29463;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Trypsin, anionic precursor (EC 3.4.21.4).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_TaxID=9913;
RP SEQUENCE FROM N.A.
RC STRAIN:Holstein-Priesian; TISSUE:Pancreas;
RX MEDLINE=91065183; PubMed=1701147;
RA le Huetou J., Micker C., Guilloreau P., Toulllec R., Puigserver A.;
RT "Isolation and nucleotide sequence of cDNA clone for bovine
RT pancreatic anionic trypsinogen. Structural identity within the
RT trypsin family.";
RL Eur. J. Biochem. 193:767-773(1990).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the Embl outstation -
CC the European Bioinformatics Institute. This statement is not subject to
CC use by non-profit institutions as long as its content and for commercial
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

PDB: 1AMH; 24-DEC-97.
PDB: 1SLU; 11-JUL-96.
PDB: 1SLV; 11-JUL-96.
PDB: 1SLW; 11-JUL-96.
PDB: 1SLX; 11-JUL-96.
PDB: 1DPO; 07-JUL-97.
PDB: 1TGI; 23-DEC-98.
PDB: 1TGF; 23-DEC-98.
MEROPS: S01.258-98.
MEROPS: S01.258-98.
InterPro: IPR001114; Chymotrypsin.
Pfam: PF00089; trypsin; 1.
PRINTS: P000722; CHYMOTRYPSIN.
SMART: SM00020; Tryp_Spc; 1.
PROSITE: PS00240; TRYPSIN_DOM; 1.
PROSITE: PS00134; TRYPSIN_HIS; 1.
PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
Multigene family; 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT SITE 84 84
FT CONFLICT 88 88
FT CONFLICT 88 88
SQ SEQUENCE 246 AA; 26228 MW; A8D3630809AE8606 CRC64;

Query Match 84.1%; Score 90; DB 1; Length 246;
Best Local Similarity 85.0%; Pred. No. 6.5e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPYQVSLNS 20
DB 24 IVGGYTCGANTVPYQVSLNS 43

RESULT 6
ID TRY2 CANFA STANDARD; PRT; 247 AA.
AC P06972;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last annotation update)
DE Trypsin, anionic precursor (EC 3.4.21.4).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]_TaxID=9615;
RP SEQUENCE FROM N.A.
RX MEDLINE=86284628; PubMed=3841794;
RA Pinsky S.D., Laforce K.S., Scheele G.;
RT "Differential regulation of trypsinogen mRNA translation: full-length
RT mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
RT in the dog pancreas.";
RL Mol. Cell. Biol. 5:2669-2676(1985).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the Embl outstation -
CC the European Bioinformatics Institute. This statement is not subject to
CC use by non-profit institutions as long as its content and for commercial
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: M15589; AAA30899.1; -.
DR PIR: A26273; TRDG.
DR HSSP: P00763; IDPO.
DR MEROPS: S01:258; -.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00200; Tryp_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 247
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT SITE 247 AA; 26423 MW; 37459031D6DBEAF CRC64;
SQ SEQUENCE 247 AA; 26423 MW; 37459031D6DBEAF CRC64;

Query Match 81.3%; Score 87; DB 1; Length 247;
Best Local Similarity 80.0%; Pred. No. 1.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 24 IVGGYTCGNSVPYQVSLNA 43

RESULT 7
TRY3_HUMAN STANDARD; PRT; 247 AA.
ID TRY3_HUMAN
AC P13951; 1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DR 15-JUN-2003 (Rel. 41, Last annotation update)
DE Trypsin III precursor (EC 3.4.21.4).
GN PRG39 OP Tryp_Precursor (EC 3.4.21.4).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RX MEDLINE=90221895; PubMed=2326201;
RA Tani T., Kawashima I., Mita K., Takiguchi Y.;
RL "Nucleotide sequence of the human pancreatic trypsinogen III cDNA.";
RL Nucleic Acids Res. 18:1631-1631(1990).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15505; CAA33527.1; -.
DR PIR: S12764; S12764.

```

```

DR HSSP: P00761; LEPT.
DR MEROPS: S01:174; -.
DR Genew: HGNC:9486; PRSS3.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00200; Tryp_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; FALSE NEG.
DR Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 247
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 194 194
FT SITE 247 AA; 26776 MW; 697DE163FCEBD06 CRC64;
SQ SEQUENCE 247 AA; 26776 MW; 697DE163FCEBD06 CRC64;

Query Match 81.3%; Score 87; DB 1; Length 247;
Best Local Similarity 80.0%; Pred. No. 1.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 24 IVGGYTCGNSVPYQVSLNS 43

RESULT 8
TRY4_HUMAN STANDARD; PRT; 304 AA.
ID TRY4_HUMAN
AC P35030; Q15665; OSUOV3.
DT 01-SEP-1994 (Rel. 28, Created)
DR 01-SEP-1994 (Rel. 28, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin IV precursor (EC 3.4.21.4) (Brain trypsinogen)
GN PRSS4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]_TaxID=9606;
RN SEQUENCE FROM N.A. (ISOFORMS A AND B).
RP TISSUE=Brain;
RX MEDLINE=94123994; PubMed=8294000;
RA Wiesand U., Corbach S., Minn A., Kang J., Mueller-Hill B.;
RT "Cloning of the cDNA encoding human brain trypsinogen and
RT characterization of its product.";
RL Gene 136:167-175(1993).
[2]
RN SEQUENCE FROM N.A. (ISOFORM C).
RP Fukuoka S.;
RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; A (shown here), B and C; seem
CC to be produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: PANCREAS AND BRAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15505; CAA33527.1; -.
DR PIR: S12764; S12764.

```


CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; X72761; CAB58178.1; -
DR EMBL; X71345; CA50484.1; -
DR EMBL; X54177; BAA0257.1; -
DR EMBL; X54177; BAA0257.1; -
DR HSP; P07496; S33496.
DR MEROPS; S01.174; ITRN.
DR InterPro; IP0001314; Chymotrypsin.
DR InterPro; IP0001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family; Alternative splicing.
FT SIGNAL 1 ?
FT PROPEP 7 80
FT CHAIN 81 304
FT ACT_SITE 120 120
FT ACT_SITE 164 164
FT ACT_SITE 287 287
FT DISULFID 105 121
FT DISULFID 196 243
FT DISULFID 228 242
FT DISULFID 253 277
FT SITE 251 251
FT VARSPIC 1 45
FT VARSPIC 1 70
FT CONFLICT 89 89
FT SEQUENCE 304 AA; 32499 MW; 4C4316C31F1D0FFC CRC64;
Query Match 81.3%; Score 87; DB 1; Length 304;
Residues Similarity 80.0%; Pred. No. 2.3e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 IVGGVTCGANTVPQVSLNS 20
Db 81 IVGGVTCGANTVPQVSLNS 100
RESULT 9
ID TRY1_RAT STANDARD; PRT: 246 AA.
AC P00762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
OS Trypsin 1, anionic precursor (EC 3.4.21.4) (pretrypsinogen I).
OC Euteleostomi (Rat).
OC Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eumetazoa; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP STRAIN-Sprague-Dawley;
RC MEDLINE=92265624; PubMed=6896710;
RA McDonald R.J., Stary S.J., Swift G.H.;
RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
RL sequences of the cloned cDNAs.";
RN [2]
RN J. Biol. Chem. 257:9724-9732(1982).
RP SEQUENCE FROM N.A.
RA MEDLINE=45054880; PubMed=6094547;
RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,

```

```

RA Rutter W.J.;
RA Structure of two related rat pancreatic trypsin genes.";
RA J. Biol. Chem. 259:14255-14264(1984).
RP [3]
RP X-RAY CRYSTALLOGRAPHY (3.8 ANGSTROMS).
RN MEDLINE=97292121; PubMed=912742;
RX Spang S., Standing T., Fletcher R.J., Stroud R.M., Finer-Moore J.,
RX Xiong N.-H., Hamlin R., Rutter W.J., Craik C.S.;
RA "The three-dimensional structure of Asn102 mutant of trypsin: role of
RT Asn102 in serine protease catalysis.";
RL Science 237:905-909(1987).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- MISCELLANEOUS: THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR
CC FORM OF TRYPSIN PRODUCED BY THE ADULT PANCREAS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by any of the participating institutions as long as its content is in no way
CC modified and this statement is included in the header of the entry.
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01273; CA24580.1; -
DR EMBL; J00778; AAA98518.1; -
DR PIR; A00948; TRPT1.
DR PDB; ITRM; 15-JUL-93.
DR PDB; 2TRM; 16-JUL-93.
DR PDB; 1BBA; 30-APR-94.
DR PDB; 1BBR; 31-JUL-94.
DR PDB; 1BRC; 31-MAY-94.
DR MEROPS; S01.094; -
DR InterPro; IP0001314; Chymotrypsin.
DR InterPro; IP0001254; Ser_protease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family; 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 179 206
FT DISULFID 171 255
FT DISULFID 196 243
FT SITE 194 194
FT SITE 251 251
FT STRAND 28 29
FT TURN 32 33
FT TURN 36 37
FT STRAND 38 42
FT STRAND 46 54
FT TURN 55 56
FT STRAND 57 60
FT HELIX 62 64
FT STRAND 70 73
FT STRAND 77 77
FT TURN 78 79
FT STRAND 86 95
FT TURN 97 98
FT STRAND 100 100
FT TURN 101 104
REQUIRED FOR SPECIFICITY.

```

```

FT STRAND 105 105
FT TURN 106 106
FT STRAND 109 113
FT STRAND 127 126
FT TURN 126 126
FT STRAND 135 126
FT TURN 135 126
FT STRAND 150 165
FT TURN 150 165
FT HELIX 168 174
FT TURN 174 174
FT TURN 176 178
FT TURN 181 182
FT STRAND 183 186
FT TURN 183 186
FT TURN 189 190
FT STRAND 194 194
FT TURN 197 198
FT STRAND 200 201
FT STRAND 203 206
FT TURN 207 208
FT STRAND 209 216
FT TURN 223 224
FT STRAND 227 231
FT STRAND 232 234
FT HELIX 236 244
FT TURN 245 245
SQ SEQUENCE 246 AA; 25959 MW; 6AFA0DAD11943FB5 CRC64;

Query Match 79.4%; Score 85; DB 1; Length 246;
Best Local Similarity 80.0%; Pred. No. 3.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPVQVSLNS 20
Db 24 IVGGYTCPEHSVPVQVSLNS 43

||||| :|||
||||| :|||

RESULT 10
TRY1 HUMAN
ID TRY1_HUMAN STANDARD; PRT; 247 AA.
AC P07477; O92955; O9HAN4; OSHANS; OSHANG; O9HAN7;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin 1 precursor (EC 3.4.21.4) (Cationic trypsinogen).
GN PRSSI OR TRY1 OR TRY1 OR TRY1 OR TRY1.
OS Homo sapiens (Human)
OC Chordata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS 1;
SC SEQUENCE FROM N.A.
RX MEDLINE=96221712; PubMed=3011602;
RA Kim M., Nakamura Y., Ogawa M., Yamamoto T., Mori T.,
RT cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsinogens;
RL Gene 41:305-310(1986).
RP [2]
RD SEQUENCE OF 16-43.
RX MEDLINE=90091010; PubMed=2598466;
RA Kimland M., Ruesick C., Marks W.H., Borgstrom A.;
RT Immunoreactive anionic and cationic trypsin in human serum;
RL Clin. Chim. Acta 184:31-46(1989).
RP [3]
RD SEQUENCE OF 68-151 FROM N.A., AND VARIANT HPC HIS-122.
RX MEDLINE=96438847; PubMed=8841182;
RA Whitcomb D.C., Gorry M.C., Preston R.A., Furey W., Sosenheimer M.J.,
RA Ulrich C.D., Martin S.P., Gates L.K. Jr., Amann S.T., Toakes P.P.,
RA Liddle R., McGrath K., Uomo G., Post J.C., Ehrlich G.D.;
RT "Hereditary pancreatitis is caused by a mutation in the cationic
RT trypsinogen gene.";
RL Nat. Genet. 14:141-145(1996).
RP [4]
RD SEQUENCE OF 15-67 FROM N.A., AND VARIANT HPC GLY-22.

```

```

RX MEDLINE=20389982; PubMed=10930381;
RA Teich N., Ockenga J., Hoffmeister A., Manna M., Mosener J., Keim V.;
RT "Chronic pancreatitis associated with an activation peptide mutation
RT that facilitates trypsin activation.";
RL Gastroenterology 119:461-465(2000).
RN [5]
RD SEQUENCE OF 68-151 FROM N.A., AND VARIANTS HPC P-104; C-116 AND F-139.
RX Teich N., Bauer N., Mosener J., Keim V.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RD X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND PHOSPHORYLATION.
RX MEDLINE=96266496; PubMed=8693601;
RA Gaboriaud C., Serre L., Guy-Crotte O., Forest E.,
RA Fontecilla-Camps J.-C.;
RT "Crystal structure of human trypsin 1: unexpected phosphorylation of
RT Tyr151.";
RL J. Mol. Biol. 259:995-1010(1996).
RN [7]
RD VARIANTS HPC ILE-29 AND HIS-122.
RX MEDLINE=97463797; PubMed=9322498;
RA Gorry M.C., Gabaizelden J., Frey W., Gates L.K. Jr., Preston R.A.,
RA Ascon C.E., Zhang L., Ulrich C., Ehrlich G.D., Whitcomb D.C.;
RT "Mutations in the cationic trypsinogen gene are associated with
RT recurrent acute and chronic pancreatitis.";
RL Gastroenterology 113:1063-1068(1997).
RN [8]
RD VARIANT HPC ILE-29.
RX MEDLINE=96295575; PubMed=9633818;
RA Teich N., Mosener J., Keim V.;
RT "Mutations of the cationic trypsinogen in hereditary pancreatitis.";
RL Hum. Mutat. 12:39-43(1998).
RN [9]
RD VARIANTS HPC VAL-16 AND HIS-122.
RX MEDLINE=9315544; PubMed=10381903;
RA Witt H., Luck W., Becker M.;
RT "A signal peptide cleavage site mutation in the cationic trypsinogen
RT gene is strongly associated with chronic pancreatitis.";
RL Gastroenterology 117:7-10(1999).
RN [10]
RD VARIANT HPC ARG-23.
RX MEDLINE=99219545; PubMed=10204851;
RA Ferec C., Raguenaes O., Salomon R., Roche C., Bernard J.P., Guillot M.,
RA Quere I., Faure C., Mercier B., Audrezet M.P., Guillausseau P.J.,
RA Dupont C., Munnich A., Bignon J.D., Le Bodic L.;
RT "Mutations in the cationic trypsinogen gene and evidence for genetic
RT heterogeneity in hereditary pancreatitis.";
RL J. Med. Genet. 36:228-232(1999).
RN [11]
RD VARIANTS HPC THR-29 AND CYS-122.
RX MEDLINE=21648565; PubMed=11788572;
RA Pfutzer R., Myers E., Applebaum-Shapiro S., Finch R., Ellis I.,
RA Neoptolemos J., Kant J.A., Whitcomb D.C.;
RT "Novel cationic trypsinogen (PRSSI) N29T and R122C mutations cause
RT autosomal dominant hereditary pancreatitis.";
RL Gut 50:271-272(2002).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MASS SPECTROMETRY: MW=24348; MW_ERR=2; METHOD=Electrospray;
CC RANGE=24-247.
CC -1- DISEASES: Defects in PRSSI are a cause of hereditary pancreatitis,
CC (HPC or HP); also known as chronic pancreatitis (CP). HPC is an
CC autosomal dominant disease characterized by the presence of
CC calculi in pancreatic ducts. It causes severe abdominal pain
CC attacks.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/announce/
CC or send an email to license@ebi.ac.uk).

```

QY		1	IUGVTCGNTGYNTPYQSLSNS 20	
DB		24	IUGVNCSESVIPYQVSLSNS 43	
<hr/>				
	RESULT 11			
ID	TRY3_CHICK	STANDARD;	PRT:	248 AA.
QY	Q90629;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
OS	Trypsin II-P29 precursor (EC 3.4.21.4).			
OS	Gallus gallus (Chicken)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Aves; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus			
OX	NCBI_TaxID=9031;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pancreas;			
RX	MEDLINE=95251611; PubMed=7733985;			
RX	Wang K., Gan L., Lee I., Hood L.E.;			
RT	"Isolation and characterization of the chicken trypsinogen gene family.";			
RT	Biochem. J. 307:471-479(1995).			
RL	-!- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -Xaa, Lys- -Xaa.			
CC	-!- SUBCELLULAR LOCATION: Extracellular.			
CC	-!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE			
CC	LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	THIS SWISS-PROT entry is copyright - it is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/ or send an email to licenses@sib-sch.ch).			
CC	EMBL; U15157; AAA79914.1; ..			
DR	HSPSP; P00763; LDPO.			
DR	MEROPS; S01.151;			
DR	InterPro; IPRO01314; Chymotrypsin.			
DR	InterPro; IPRO01254; Ser_protease_Try.			
DR	Fam; PF00089; trypsin_1.			
DR	PRINIS; PR00722; CHYMOTRYSIN.			
DR	PROSITE; PS00344; TRYPSIN_DOM; 1.			
DR	PROSITE; PS0134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;			
KW	Multigene family.			
FT	SIGNAL	1	16	BY SIMILARITY.
FT	PROPEP	17	25	ACTIVATION PEPTIDE (BY SIMILARITY).
FT	CHAIN	26	248	TRYPSIN II-P29.
FT	ACT SITE	65	65	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT SITE	109	109	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT SITE	202	202	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DIGUFID	32	162	BY SIMILARITY.
FT	DIGUFID	50	66	BY SIMILARITY.
FT	DIGUFID	134	235	BY SIMILARITY.
FT	DIGUFID	141	208	BY SIMILARITY.
FT	DIGUFID	173	187	BY SIMILARITY.
FT	DIGUFID	186	196	BY SIMILARITY.
FT	SITE	136	191	REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SEQ	SEQUENCE	248 AA;	26622 MW;	ESEI6807622B588E CRC64;
<hr/>				
	Query Match		Score 85;	DB 1; Length 248;
	Best Local Similarity		80.0%;	Pred. No. 3,9e-06;
	Matches	16;	Conservative	2; Mismatches 2; Indels 0; Gaps 0;

```

Db 26 IVGGYTCDEHSVPYQVSLNS 45
||||| :|||||
RESULT 12
ID TRV2_KENLA STANDARD; PRT; 244 AA.
AC P70059;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4)
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8621712; PubMed=3011602;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Mori T.,
RA Matsubara K.;
RT "Cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsinogens.";
RL Gene 41:305-310(1986).
RN [2]
RP SEQUENCE OF 16-49.
RX MEDLINE=90091010; PubMed=2598466;
RA Kimmelman M., Ruesick C., Marks W.H., Borgstrom A.;
RT "Immunoreactive anionic and cationic trypsin in human serum.";
RL Clin. Chim. Acta 184:31-46(1989).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC EMBL: M37602; AB017274.1; -
CC HSPG: P00763; IDPO.
CC MEROPS: S01.258; -
CC MIM: 601564; -
CC InterPro: IPR001314; Chymotrypsin.
CC Pfam: PF00089; trypsin; I.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC KMW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
CC Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 21 BY SIMILARITY.
FT CHAIN 22 244 ACTIVATION PEPTIDE (BY SIMILARITY).
FT ACT_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 28 158 BY SIMILARITY.
FT DISULFID 46 62 BY SIMILARITY.
FT DISULFID 130 231 BY SIMILARITY.
FT DISULFID 137 204 BY SIMILARITY.
FT DISULFID 169 183 BY SIMILARITY.
FT DISULFID 194 218 BY SIMILARITY.
FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 244 AA; 26079 MW; C632P3CB3300B323 CRC64;

Query Match 78.5%; Score 84; DB 1; Length 244;
Best Local Similarity 75.0%; Pred. No. 5.5e-06;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
||||| :|||||
Db 22 IVGGFTCAKNAVYQVSLNA 41
||||| :|||||
RESULT 13
ID TRV2_HUMAN STANDARD; PRT; 247 AA.
AC P07478;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II precursor (EC 3.4.21.4) (Anionic trypsinogen).
GN PRSS2 OR TRY2 OR TRY2P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8621712; PubMed=3011602;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Mori T.,
RA Matsubara K.;
RT "Cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsinogens.";
RL Gene 41:305-310(1986).
RN [2]
RP SEQUENCE OF 16-49.
RX MEDLINE=90091010; PubMed=2598466;
RA Kimmelman M., Ruesick C., Marks W.H., Borgstrom A.;
RT "Immunoreactive anionic and cationic trypsin in human serum.";
RL Clin. Chim. Acta 184:31-46(1989).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC EMBL: M37602; AB017274.1; -
CC HSPG: P00763; IDPO.
CC MEROPS: S01.258; -
CC MIM: 601564; -
CC InterPro: IPR001314; Chymotrypsin.
CC Pfam: PF00089; trypsin; I.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC KMW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
CC Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23 ACTIVATION PEPTIDE.
FT CHAIN 24 247 TRYPSIN II.
FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 30 160 BY SIMILARITY.
FT DISULFID 48 64 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 196 220 BY SIMILARITY.
FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 247 AA; 26488 MW; 8280FA41EBE3D5DB CRC64;

Query Match 78.5%; Score 84; DB 1; Length 247;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
||||| :|||||
Db 24 IVGGYTCSENSVPYQVSLNS 43
||||| :|||||

```

RESULT 14

```

ID   TRY3_RAT   STANDARD:      PRT;   247 AA.
AC   P08426;1988 (Rel. 08, Created)
DT   01-AUG-1998 (Rel. 08, Last sequence update)
DE   15-JUN-2002 (Rel. 41, Last annotation update)
DE   TRYPSIN III, cationic precursor (EC 3.4.21.4) (Pretrypsinogen III).
GN   TRY3
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RX   MEDLINE=8721609; PubMed=3607011;
RA   Fletcher T.S., Alhadef M., Craik C.S., Largman C.;
RT   Isolation and characterization of a cDNA encoding rat cationic
RT   trypsinogen. J Biol Chem 268:1081-1086 (1993).
RL   Biochemistry, 32:1081-1086 (1993).
CC   -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC   -1- SUBCELLULAR LOCATION: Extracellular.
CC   -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M16624; AAA41985.1; -.
DR   PIR; A27547; A27547.
DR   HSPG; P00763; IDPO.
DR   MROS; S01056; S114. Chymotrypsin.
DR   InterPro; IPR001314; Chymotrypsin.
DR   Pfam; PR00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; TRY_SPC; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Multigene family.
KW   Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT   SIGNAL      1..15
FT   PROPEP      16..24
FT   CHAIN        25..247
FT   ACT_SITE     64..64
FT   ACT_SITE     108..108
FT   DISULFID     21..161
FT   DISULFID     49..165
FT   DISULFID     133..234
FT   DISULFID     140..207
FT   DISULFID     172..186
FT   DISULFID     197..221
FT   SITE         195..195
SQ   SEQUENCE 247 AA; 26269 MW; D74892BAA584E4A8 CRC64;

```

```

Query Match      78.5%; Score 84; DB 1; Length 247;
Best Local Similarity 75.0%; Pred. No. 5.6e-06;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVGGTTCGANTVPYQVSLNS 20
    |||||  |||||
DB 25 IVGGTTCGNTSPYQVSLNA 44

```

RESULT 15

TRY2_MOUSE

```

ID   TRY2_MOUSE  STANDARD:      PRT;   246 AA.
AC   P0146;1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DE   15-JUN-2002 (Rel. 41, Last annotation update)
DE   TRYPSIN II, anionic precursor (EC 3.4.21.4) (Pretrypsinogen II).
GN   TRY2
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RX   MEDLINE=87066713; PubMed=3641189;
RA   Stevenson B.J., Hagenbuchle O., Wellauer P.K.;
RT   Sequence organisation and transcriptional regulation of the mouse
RT   trypsinogen gene. Nucleic Acids Res 14:8307-8330 (1986).
RL   Nucleic Acid Res 14:8307-8330 (1986).
CC   -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC   -1- SUBCELLULAR LOCATION: Extracellular.
CC   -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X04574; CAA28243.1; -.
DR   EMBL; X04577; CAA28245.1; -.
DR   HSPG; B25056; B25056.
DR   MROS; P00763; IDPO.
DR   MGD; MG1102759; Try2.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Ser. protease_Try.
DR   Pfam; PR00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; TRY_SPC; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT   SIGNAL      1..15
FT   PROPEP      16..23
FT   CHAIN        24..246
FT   ACT_SITE     64..64
FT   ACT_SITE     107..107
FT   ACT_SITE     200..200
FT   DISULFID     30..160
FT   DISULFID     48..64
FT   DISULFID     132..233
FT   DISULFID     139..206
FT   DISULFID     171..185
FT   DISULFID     196..220
FT   SITE         194..194
SQ   SEQUENCE 246 AA; 26203 MW; CEF8C97AACD07AD CRC64;

```

```

Query Match      76.6%; Score 82; DB 1; Length 246;
Best Local Similarity 75.0%; Pred. No. 1.1e-05;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVGGTTCGANTVPYQVSLNS 20
    |||||  |||||
DB 24 IVGGTTCRESSVPYQVSLNA 43

```

```

Search completed: February 12, 2003, 10:23:12
Job time : 2.68657 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:16:06 ; Search time 14.1493 Seconds
(without alignments)
291.248 Million cell updates/sec

Title: US-10-036-371-2

Perfect score: 107
Sequence: 1 IGVGYTCGANTVPYQVSLNS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_invertebrate:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mcy:*
8: sp_organella:*
9: sp_phase:*
10: sp_plant:*
11: sp_rhodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_tvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	92	86.0	24 11 Q9QV67	Q9QV67 rattus sp.
2	90	84.1	246 11 Q9ZLR9	Q9ZLR9 mus musculus
3	90	84.1	246 11 Q9R0T7	Q9R0T7 mus musculus
4	89	83.2	246 11 Q9QUK9	Q9QUK9 mus musculus
5	86	80.4	247 11 Q9CPN9	Q9CPN9 mus musculus
6	82	76.6	247 11 Q9D7Y7	Q9D7Y7 mus musculus
7	81	75.7	247 11 Q9CPN7	Q9CPN7 mus musculus
8	76	71.0	237 13 Q91515	Q91515 fugu rubrip
9	71	66.4	244 13 Q42159	Q42159 petromyzon
10	71	66.4	245 13 Q42160	Q42160 petromyzon
11	67	62.6	240 13 Q9RH0R	Q9RH0R engraulis j
12	66	61.7	247 13 Q9WQ55	Q9WQ55 paratichthys
13	65	60.7	248 13 Q9B7G9	Q9B7G9 paratichthys
14	65	60.7	248 13 Q9B7G9	Q9B7G9 paratichthys
15	65	60.7	242 13 Q9K7Q7	Q9K7Q7 paratichthys
16	62	57.9	247 13 Q42158	Q42158 petromyzon

17	62	57.9	247 13 Q42608	Q42608 petromyzon
18	61	57.0	244 13 Q9QCM3	Q9QCM3 anguilla ja
19	61	57.0	249 13 Q9M6K0	Q9M6K0 nototheria
20	60	56.1	117 13 Q9DFJ6	Q9DFJ6 gillichthys
21	60	56.1	219 13 Q91036	Q91036 gadus morhua
22	60	56.1	242 13 Q93266	Q93266 pseudopleur
23	60	56.1	249 13 Q92046	Q92046 disostichu
24	60	56.1	250 13 Q93265	Q93265 pseudopleur
25	60	56.1	250 13 Q93265	Q93265 disostichu
26	59	55.1	344 13 Q9W6J9	Q9W6J9 disostichu
27	58	54.2	242 13 Q92099	Q92099 petromotile
28	57	53.3	17 11 Q9QVC2	Q9QVC2 rattus sp.
29	57	53.3	248 11 Q9QV77	Q9QV77 rattus sp.
30	57	53.3	248 11 Q9QV77	Q9QV77 rattus sp.
31	57	53.3	248 11 Q9QV77	Q9QV77 rattus sp.
32	56	52.3	248 11 Q9QV77	Q9QV77 rattus sp.
33	56	52.3	248 11 Q9QV77	Q9QV77 rattus sp.
34	55	51.4	245 5 Q9XV47	Q9XV47 ctenosopha
35	55	51.4	245 5 Q9XV47	Q9XV47 ctenosopha
36	55	51.4	245 5 Q9XV47	Q9XV47 ctenosopha
37	53	49.5	261 11 Q88109	Q88109 mus musculus
38	52	48.6	239 11 Q63275	Q63275 rattus norv
39	52	48.6	239 11 Q63275	Q63275 rattus norv
40	52	48.6	239 11 Q63275	Q63275 rattus norv
41	52	48.6	246 5 Q9XY46	Q9XY46 ctenosopha
42	52	48.6	505 16 Q98NR7	Q98NR7 thizobolium l
43	51	47.7	180 5 Q9U7D2	Q9U7D2 caenorhabdi
44	51	47.7	261 5 Q962G7	Q962G7 culex pipie
45	51	47.7	266 5 Q24091	Q24091 drosophila

ALIGNMENTS

RESULT 1	Q9QV67	PRELIMINARY;	PRT;	24 AA.
AC	Q9QV67;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Caltrin (Fragment).			
OS	Rattus sp.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
RN	NCBI_TaxID=10116;			
RN	SEQUENCE			
RP	MEDLINE:330508; PubMed:8118586;			
RX	Medline:330508; PubMed:8118586;			
RT	Caltrin, the calcium transport inhibitor protein from seminal vesicles of the rat;			
RL	Biol. Reprod. 48:1326-1333(1993).			
RQ	HSP: P00760; 1A07.			
SQ	SEQUENCE 24 AA; 2452 MW; DD1A6C82EF25FC1A CRC64;			
Query Match	86.0%;	Score 92;	DB 11;	Length 24;
Best Local Similarity	94.7%;	Pred. No. 1e-07;		
Matches 18;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	2 VGGYTCGANTVPYQVSLNS 20			
DB	2 VGGYTCGANTVPYQVSLNS 20			
RESULT 2	Q9ZLR9	PRELIMINARY;	PRT;	246 AA.
AC	Q9ZLR9;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			


```

RC TISSUE=TESTIS;
RX MEDLINE=9436155; PubMed=10506205;
RA Omura K., Konno N., Kobayashi Y., Yamagata K., Sato S.,
RT "Acrosome of bacnetric trypsin is localized in the acrosome of
RT mammalian sperm and is released during acrosome reaction.";
RL J Biol. Chem. 274:29426-29432(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiri L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=21103195; PubMed=11160223;
RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT "Differential transcriptional regulation of individual TCR Vbeta
RT segments before gene rearrangement.";
RL J Immunol. 166:1771-1780(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB009661; BA85187.1; -.
DR EMBL; AB017031; BA474760.1; -.
DR EMBL; AK007843; BAB25300.1; -.
DR EMBL; AE000664; AAB69057.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.058; -.
DR MED; MG1.1913350; 0910001B191K.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TYP_SPC.1.
DR PROSITE; PS00240; TRYPsin DOM.1.
DR PROSITE; PS00134; TRYPsin HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin SER.1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 246 AA; 26277 MW; 915C92353EC42809 CRC64;

Query Match 83.2%; Score 89; DB 11; Length 246;
Best Local Similarity 80.0%; Pred. No. 3.9e-06;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPQVSLNS 20
DB 24 IVGGYTCRENSIFQVSLNS 43

RESULT 5
QY QYCGNG PRELIMINARY; PRT; 247 AA.
AC QYCGNG;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

Query Match 80.4%; Score 86; DB 11; Length 247;
Best Local Similarity 80.0%; Pred. No. 1.2e-05;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYTCGANTVPQVSLNS 20
DB 25 IVGGYTCQNALPQVSLNS 44

RESULT 6
QY QY7Y7 PRELIMINARY; PRT; 247 AA.
AC QY7Y7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE 2210010C04RIK protein.
 GN 2210010C04RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saïto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saïto R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Maesano Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scabuli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Flegler N., Carninci P., de Bonaldo M.F.,
 RA Guscinich S., Hill D., Flegler N., Carninci P., Fujita M., Gariboldi M.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690 (2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AK008695; BAB25837.1; -;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.151; -;
 DR MGD: MGI:1914623; 2210010C04RIK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc. 1.
 DR PROSITE: PS02040; TRYPSIN_DOM. 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 247 AA; 26407 MW; 8AD47DB32283A55 CRC64;
 Query Match 76.6%; Score 82; DB 11; Length 247;
 Best Local Similarity 78.9%; Pred. No. 5e-05;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 VGGYTCGANTVPQVSLNS 20
 Db 26 VGGYTCORNALPQVSLNS 44
 ||||| |
 RESULT 7
 Q9C9N7 PRELIMINARY; PRT; 247 AA.
 ID Q9C9N7
 AC Q9C9N7
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DE 2210010C04RIK protein.
 GN 2210010C04RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saïto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saïto R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Maesano Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scabuli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Flegler N., Carninci P., de Bonaldo M.F.,
 RA Guscinich S., Hill D., Flegler N., Carninci P., Fujita M., Gariboldi M.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690 (2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AK008695; BAB25837.1; -;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.151; -;
 DR MGD: MGI:1914623; 2210010C04RIK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc. 1.
 DR PROSITE: PS02040; TRYPSIN_DOM. 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 247 AA; 26407 MW; 8AD47DB32283A55 CRC64;
 Query Match 76.6%; Score 82; DB 11; Length 247;
 Best Local Similarity 78.9%; Pred. No. 5e-05;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 VGGYTCGANTVPQVSLNS 20
 Db 26 VGGYTCORNALPQVSLNS 44
 ||||| |
 RESULT 7
 Q9C9N7 PRELIMINARY; PRT; 247 AA.
 ID Q9C9N7
 AC Q9C9N7
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DE 2210010C04RIK protein.
 GN 2210010C04RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saïto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saïto R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Maesano Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scabuli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Flegler N., Carninci P., de Bonaldo M.F.,
 RA Guscinich S., Hill D., Flegler N., Carninci P., Fujita M., Gariboldi M.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690 (2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AK007406; BAB25018.1; -;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.129; -;
 DR MGD: MGI:1920876; 1810009J06RIK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc. 1.
 DR PROSITE: PS02040; TRYPSIN_DOM. 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 247 AA; 26503 MW; EDF5F7696833C7BC CRC64;
 Query Match 75.7%; Score 81; DB 11; Length 247;
 Best Local Similarity 78.9%; Pred. No. 7.2e-05;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 IUGGYTCGANTVPQVSLN 19
 Db 24 IUGGYTCFRESVPQVSLN 42
 ||||| |
 RESULT 8
 Q91515 PRELIMINARY; PRT; 237 AA.
 ID Q91515
 AC Q91515
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Trypsinogen (Fragment).
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OC NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saïto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saïto R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Maesano Y., Nikaïdo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Scabuli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Flegler N., Carninci P., de Bonaldo M.F.,
 RA Guscinich S., Hill D., Flegler N., Carninci P., Fujita M., Gariboldi M.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690 (2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL: AK007406; BAB25018.1; -;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.129; -;
 DR MGD: MGI:1920876; 1810009J06RIK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc. 1.
 DR PROSITE: PS02040; TRYPSIN_DOM. 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 247 AA; 26503 MW; EDF5F7696833C7BC CRC64;
 Query Match 75.7%; Score 81; DB 11; Length 247;
 Best Local Similarity 78.9%; Pred. No. 7.2e-05;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 IUGGYTCGANTVPQVSLN 19
 Db 24 IUGGYTCFRESVPQVSLN 42
 ||||| |
 RESULT 8
 Q91515 PRELIMINARY; PRT; 237 AA.
 ID Q91515
 AC Q91515
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Trypsinogen (Fragment).
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OC NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RA Wang K., Gan L., Lee I., Roach J., Hood J.,
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: U25747; AAA75001.1; -.
DR HSP: P35031; 1BIT.
DR MEROPS: S01.128; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PSS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR HYDROLASE: Serine protease; Signal.
FT NON_TER <1 1
FT CHAIN 14 244
FT SEQUENCE 237 AA; 25726 MW; 30D2DBAC39080C2 CRC64;
SQ
Query Match 71.0%; Score 76; DB 13; Length 237;
Best Local Similarity 75.0%; Pred. No. 0.0042;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLN 20
Db 16 IVGGYECRNSVAYQVSLN 35

RESULT 9
042159
ID O42159 PRELIMINARY; PRT; 244 AA.
AC O42159
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen B2 precursor (Fragment).
GN TRYPSIN.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RT "The Molecular Evolution of the Vertebrate Trypsinogens.";
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AF011901; AAB69656.1; -.
DR HSP: P00763; 1DPO.
DR MEROPS: S01.128; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PSS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR HYDROLASE: Serine protease; Signal.
FT NON_TER <1 1
FT CHAIN 14 245
FT SEQUENCE 245 AA; 26001 MW; 9A912508B896C93E CRC64;
SQ
Query Match 66.4%; Score 71; DB 13; Length 245;
Best Local Similarity 68.4%; Pred. No. 0.0027;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLN 19
Db 22 IVGGYECARHSQPMQVSLN 40

RESULT 11
Q98TH0
ID Q98TH0 PRELIMINARY; PRT; 240 AA.
AC Q98TH0
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OX Engraulis.
OX NCBI_TaxID=42892;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Matabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA.";
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AB04329; BAB40329.1; -.
DR HSP: P00763; 1DPO.
DR MEROPS: S01.128; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.

```

```

DR Pfam: PF00089; trypsin: 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 240 AA; 26026 MW; 299B19BFF071464 CRC64;

Query Match
Best Local Similarity 62.6%; Score 67; DB 13; Length 240;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 20 IVGGYTCGANTVPQVSLNS 39

RESULT 12
QSWTQS PRELIMINARY; PRT; 247 AA.
AC QSWTQS, 1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen 3.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes;
OC Pleuronectidae; Paralichthyidae; Paralichthys.
CX NCBI_TaxID=8255;
RN 1
RP TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for trypsinogen 3.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
EMBL: AB029752; BAA82364.2; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.124; -.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin: 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 247 AA; 26948 MW; DC4B647179DD972 CRC64;

Query Match
Best Local Similarity 61.7%; Score 66; DB 13; Length 247;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPQVSLN 19
Db 22 IVGGYTCGANTVPQVSLN 40

RESULT 13
QSWTQS PRELIMINARY; PRT; 238 AA.
AC QSWTQS, 1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen 2 (Fragment).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Perciformes;
OC Pleuronectidae; Paralichthyidae; Paralichthys.
CX NCBI_TaxID=8255;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for trypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
EMBL: AB029751; BAA82363.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin: 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
RN NON_TER 1
SQ SEQUENCE 238 AA; 26071 MW; F2B8908085B8D62 CRC64;

Query Match
Best Local Similarity 60.7%; Score 65; DB 13; Length 238;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 18 IVGGYTCGANTVPQVSLNS 37

RESULT 14
QSWTQS PRELIMINARY; PRT; 241 AA.
AC QSWTQS, 1999 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen II.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OC Engraulis.
CX NCBI_TaxID=42892;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
EMBL: AB041930; BAB40330.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.258; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin: 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 241 AA; 26282 MW; FE362D39CAEEB2F6 CRC64;

Query Match
Best Local Similarity 60.7%; Score 65; DB 13; Length 241;

```

Best Local Similarity 65.0%; Pred. No. 0.024; Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
 Db 20 IVGGYECOPYSQPHQVSLNS 39

RESULT 15

Q9M707 PRELIMINARY; PRT; 242 AA.
 AC Q9M7Q7;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 OS Trypsinogen 1, livivacus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidei; Paralichthyidae; Paralichthys.
 OC NCBI_TaxID=8255;
 OK NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RA Suzuki T., Srivastava A.S., Kurokawa T.;
 RT "Japanese flounder mRNA for trypsinogen 1.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AB029750; BAA82362.1; -.
 DR HSRF; P00763; IDPO.
 DR MEROPS; S01.125;
 DR Incefto; IP001314; Chymotrypsin.
 DR IPIPF; IP001254; Serineprotease_Try.
 DR PISMP; PF001089; TRYD618.1
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS0135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 242 AA; 26548 MW; 6DA722C80BC194A2 CRC64;

Query Match 60.7%; Score 65; DB 13; Length 242;
 Best Local Similarity 65.0%; Pred. No. 0.024;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
 Db 21 IVGGYECTPYSQPHQVSLNS 40

Search completed: February 12, 2003, 10:27:18
 Job time : 15.1493 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:18:40 ; Search time 4.29851 Seconds
(without alignments)
136.898 Million cell updates/sec

Title: US-10-036-371-2

Perfect score: 107

Sequence: 1 TVGGYTCGANTVPQVSLNS 20

Scoring table: BIASUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs; 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/1/ias/5a.COMB.pcp.*
- 2: /cgn2_6/prodata/1/ias/5b.COMB.pcp.*
- 3: /cgn2_6/prodata/1/ias/5a.COMB.pcp.*
- 4: /cgn2_6/prodata/1/ias/5b.COMB.pcp.*
- 5: /cgn2_6/prodata/1/ias/6b.COMB.pcp.*
- 6: /cgn2_6/prodata/1/ias/6b.COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	107	100.0	223	1	US-08-278-091-9
2	107	100.0	223	1	US-08-483-859-9
3	107	100.0	223	1	US-08-482-173-9
4	107	100.0	223	1	US-08-487-167-9
5	107	100.0	223	2	US-08-296-949-9
6	107	100.0	223	2	US-08-815-499-9
7	107	100.0	223	2	US-08-615-271-9
8	107	100.0	223	3	US-09-074-660-9
9	107	100.0	223	3	US-09-074-659-9
10	107	100.0	223	3	US-09-106-468-9
11	107	100.0	223	3	US-09-106-468-9
12	107	100.0	223	4	US-09-106-468-9
13	107	100.0	223	4	US-09-106-468-9
14	107	100.0	223	3	US-09-120-582-2
15	96	89.7	247	2	US-08-956-267A-2
16	87	81.3	224	4	US-08-944-483-35
17	85	79.4	224	2	US-08-766-982-13
18	85	79.4	224	4	US-08-944-483-36
19	85	79.4	224	4	US-08-944-483-36
20	82	79.4	225	2	US-09-027-337-5
21	82	79.4	225	2	US-09-027-337-5
22	85	78.4	226	2	US-09-644-600-5
23	84	78.5	226	2	US-08-478-4048-44
24	68.5	64.0	281	1	US-08-467-155A-7
25	68.5	64.0	281	2	US-08-628-198-7
26	68.5	64.0	281	4	US-09-201-038-7
27	68.5	64.0	281	5	PCT-US96-07343-7

Sequence 9, Appli
Sequence 9, Appli
Sequence 9, Appli
Sequence 9, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 17, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 20, Appli
Sequence 20, Appli
Sequence 20, Appli
Sequence 20, Appli
Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-08-278-091-9
; Sequence 9, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hinx7 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; FIRM: Stewart, Michael
; REGISTRATION NUMBER: 1,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-278-091-9

Query Match 100.0%; Score 107; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVGGYTCGANTVPQVSLNS 20
|||||

```

Db      1 IVGGYTCGANTVPYQVSLNS 20

RESULT 2
US-08-483-859-9
; Sequence 9, Application US/08483859
; Patent No. 5656436
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 81m & McBurney
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,859
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY INFORMATION:
; PRIORITY NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:VG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-472-173-9
; Query Match 100.0%; Score 107; DB 1; Length 223;
; Best Local Similarity 100.0%; Pred. No. 1.5e-08;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 IVGGYTCGANTVPYQVSLNS 20
Db      1 IVGGYTCGANTVPYQVSLNS 20

RESULT 4
US-08-487-167-9
; Sequence 9, Application US/08487167
; Patent No. 5869302
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 81m & McBurney
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,173
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY INFORMATION:
; PRIORITY NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:VG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-483-859-9
; Query Match 100.0%; Score 107; DB 1; Length 223;
; Best Local Similarity 100.0%; Pred. No. 1.5e-08;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 IVGGYTCGANTVPYQVSLNS 20
Db      1 IVGGYTCGANTVPYQVSLNS 20

RESULT 3
US-08-472-173-9
; Sequence 9, Application US/08472173
; Patent No. 5655353
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 81m & McBurney
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

```



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,167
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-487-167-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGGVTCGANTVPQVSLNS 20
DB 1 IVGGVTCGANTVPQVSLNS 20

RESULT 5
US-08-482-816-9
; Sequence 9, Application US/08482a16
; Patent No. 5935573
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,816
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-482-816-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGGVTCGANTVPQVSLNS 20
DB 1 IVGGVTCGANTVPQVSLNS 20

RESULT 6
US-08-296-149-9
; Sequence 9, Application US/08296149
; Patent No. 5939297
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,149
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-296-149-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGGVTCGANTVPQVSLNS 20
DB 1 IVGGVTCGANTVPQVSLNS 20

RESULT 7
US-08-296-149-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVGGVTCGANTVPQVSLNS 20
DB 1 IVGGVTCGANTVPQVSLNS 20
```

```

US-08-801-499-9
; Sequence 9, Application US/08801499
; Patent No. 5962430
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,499
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:Jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,816
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:Jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-801-499-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 1 IVGGYTCGANTVPQVSLNS 20

RESULT 8
US-08-615-271-9
; Sequence 9, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-615-271-9

Query Match 100.0%; Score 107; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 1 IVGGYTCGANTVPQVSLNS 20

RESULT 9
US-09-074-660-9
; Sequence 9, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-659-9
Query Match 100.0%; Score 107; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

RESULT 10
US-09-074-659-9
; Sequence 9, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,659
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155

```

```

; TELEFAX: (416) 595-1163
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-659-9
Query Match 100.0%; Score 107; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

RESULT 11
US-09-106-468-9
; Sequence 9, Application US/09106468
; Patent No. 6114125
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,468
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 06-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-106-468-9
Query Match 100.0%; Score 107; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

```

```

RESULT 12
US-09-106-466A-9
; Sequence 9, Application US/09106466A
; Patent No. 6147057
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,466A
; FILING DATE: 20-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-826
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-466A-9

Query Match 100.0%; Score 107; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

RESULT 13
US-09-106-467-9
; Sequence 9, Application US/09106467
; Patent No. 6153580
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,467A
; FILING DATE: 20-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-826
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-466A-9

Query Match 100.0%; Score 107; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

RESULT 14
US-09-120-582-2
; Sequence 2, Application US/09120582
; Patent No. 6087558
; GENERAL INFORMATION:
; APPLICANT: Hood, Elizabeth
; APPLICANT: Howard, John A.
; TITLE OF INVENTION: COMMERCIAL PRODUCTION OF PROTEASES IN PLANTS
; FILE REFERENCE: 10014
; CURRENT APPLICATION NUMBER: US/09/120,582
; CURRENT FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Bos taurus
; US-09-120-582-2

Query Match 100.0%; Score 107; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 7 IVGGYTCGANTVPYQVSLNS 26

RESULT 15
US-08-956-267A-2

```

```

; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,467
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-467-9

Query Match 100.0%; Score 107; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 1 IVGGYTCGANTVPYQVSLNS 20

RESULT 14
US-09-120-582-2
; Sequence 2, Application US/09120582
; Patent No. 6087558
; GENERAL INFORMATION:
; APPLICANT: Hood, Elizabeth
; APPLICANT: Howard, John A.
; TITLE OF INVENTION: COMMERCIAL PRODUCTION OF PROTEASES IN PLANTS
; FILE REFERENCE: 10014
; CURRENT APPLICATION NUMBER: US/09/120,582
; CURRENT FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Bos taurus
; US-09-120-582-2

Query Match 100.0%; Score 107; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVGGYTCGANTVPYQVSLNS 20
Db 7 IVGGYTCGANTVPYQVSLNS 26

RESULT 15
US-08-956-267A-2

```

```

; Sequence 2, Application US/08956267A
; Patent No. 5945328
; GENERAL INFORMATION:
; APPLICANT: KJELDSEN, Helle Fabricius
; APPLICANT: KJELDSEN, Thomas Borglum
; TITLE OF INVENTION: A Process For Producing Trypsin
; TITLE OF INVENTION: (Trypsinogen)
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5945328o No. 5945328disk of No. 5945328th America, Inc.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956.267A
; FILING DATE: 22-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4500.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-6523
; TELEFAX: 212-878-6525
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-956-267A-2

```

```

Query Match      89.7%; Score 96; DB 2; Length 247;
Best Local Similarity 85.0%; Pred. No. 6.9e-07;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  1  IVGGYTCGANTVPYQVSLNS 20
Db   25  IVGGYTCGANSIPYQVSLNS 44

```

```

Search completed: February 12, 2003, 10:30:06
Job time : 4.29851 secs

```

THIS PAGE BLANK (USPTO)


```

/ TITLE OF INVENTION: Structures of Molecules
/ FILE REFERENCE: 522.192102
/ CURRENT APPLICATION NUMBER: US/09/910,071
/ CURRENT FILING DATE: 2001-07-23
/ PRIOR APPLICATION NUMBER: US 08/014,867
/ PRIOR FILING DATE: 1993-02-08
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 14
/ LENGTH: 223
/ TYPE: PRT
/ ORGANISM: unknown
/ FEATURE:
/ OTHER INFORMATION: amino acid sequence of trypsin
US-09-910-071-14

Query Match          91.6%; Score 98; DB 10; Length 223;
Best Local Similarity 95.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy  1  IVGGYTCGANTVPYQVSLNS 20
      ||||| ||||| ||||| |||||
Db  1  IVGGYTCCANTVPYQVSLNS 20

RESULT 3
US-09-923-779-154
/ Sequence 154, Application US/09923779
/ Patent No. US20020076721A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Kalos, Michael D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 522.192102
/ FILE REFERENCE: 2002101353
/ CURRENT APPLICATION NUMBER: US/09/923,779
/ CURRENT FILING DATE: 2001-08-06
/ NUMBER OF SEQ ID NOS: 155
/ SOFTWARE: PsetSeq for Windows Version 4.0
/ SEQ ID NO 154
/ LENGTH: 247
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-923-779-154

Query Match          78.5%; Score 84; DB 10; Length 247;
Best Local Similarity 80.0%; Pred. No. 2.2e-05;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy  1  IVGGYTCGANTVPYQVSLNS 20
      ||||| ||||| ||||| |||||
Db  24  IVGGYICRENSVPYQVSLNS 43

RESULT 4
US-10-021-368-7
/ Sequence 7, Application US/10021368
/ Patent No. US20020106367A1
/ GENERAL INFORMATION:
/ APPLICANT: Band, Vimala
/ TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
/ MOLECULES AND METHODS
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/10/021,368
/ FILING DATE: 12-Dec-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION NUMBER: 09/201,038
/ APPLICATION DATA:
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00398/100002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ MEDIUM TYPE: Floppy disk

```

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/10/021,368
/ FILING DATE: 12-Dec-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION NUMBER: 09/201,038
/ APPLICATION DATA:
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00398/100002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 7;
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 281 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-021-368-7

Query Match          64.0%; Score 68.5; DB 12; Length 281;
Best Local Similarity 55.6%; Pred. No. 0.0049;
Matches 15; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Oy  1  IVGGY-----TCGANTVPYQVSLNS 20
      ||||| ||||| ||||| |||||
Db  24  IVGGYMTYRTYCRESSVPYQVSLNA 50

RESULT 5
US-10-021-368-9
/ Sequence 9, Application US/10021368
/ Patent No. US20020106367A1
/ GENERAL INFORMATION:
/ APPLICANT: Band, Vimala
/ TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
/ MOLECULES AND METHODS
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/10/021,368
/ FILING DATE: 12-Dec-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION NUMBER: 09/201,038
/ APPLICATION DATA:
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00398/100002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154

```


[illegible]

```

; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PR1
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (6)
; OTHER INFORMATION: Q or E
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (8)
; OTHER INFORMATION: E or T
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (9)
; OTHER INFORMATION: K or R
; US-10-036-371-1

```

```

Query Match      58.9%; Score 63; DB 12; Length 20;
Best Local Similarity 60.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 1 IVGGYCXKHQSQHQSLSNS 20

```

```

RESULT 10
US-10-036-371-8
; Sequence 8, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PR1
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (9)
; OTHER INFORMATION: K or R
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (24)
; OTHER INFORMATION: Y or F
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (12)
; OTHER INFORMATION: K or E
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (13)
; OTHER INFORMATION: D or Q
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (64)
; OTHER INFORMATION: Y or F

```

```

; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (7)
; OTHER INFORMATION: I or unknown
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (82)
; OTHER INFORMATION: N or D
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (92)
; OTHER INFORMATION: K or E
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (125)
; OTHER INFORMATION: A or D
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (136)
; OTHER INFORMATION: V or C
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (138)
; OTHER INFORMATION: N or S
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (204)
; OTHER INFORMATION: H or N
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (212)
; OTHER INFORMATION: M, V or C
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (224)
; OTHER INFORMATION: N or S
; US-10-036-371-8

```

```

Query Match      57.0%; Score 61; DB 12; Length 225;
Best Local Similarity 60.0%; Pred. No. 0.005;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 IVGGYTCGANTVPQVSLNS 20
Db 1 IVGGYCXKHQSQHQSLSNS 20

```

```

RESULT 11
US-09-938-269-15
; Sequence 15, Application US/09938269
; Publication No. US20030007951A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; APPLICANT: Cowling, Didier S.P.
; APPLICANT: Hubbel, Jeffrey A.
; APPLICANT: van de Wetering, Petra
; TITLE OF INVENTION: Treatment of Trauma
; FILE REFERENCE: 314572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 25
; TYPE: PR1
; ORGANISM: Atlantic cod
; US-09-938-269-15

```

```

Query Match      56.1%; Score 60; DB 9; Length 25;
Best Local Similarity 60.0%; Pred. No. 0.0073;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 IVGGYTCGANTVPQVSLNS 20

```

```

DB      1  IVGGYECTRHQAQVSLNS 20
      TYPE: amino acid
      STRANDEDNESS: not relevant
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 10:
      US-10-021-368-10

RESULT 12
US-09-938-269-16
; Sequence 16, Application US/09938269
; Publication No. US2003000795A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; APPLICANT: Cowling, Didier S.P.
; APPLICANT: Hubbel, Jeffrey A.
; APPLICANT: Hubbel, Jeffrey A.
; TITLE OF INVENTION: Treatment Of Trauma
; FILE REFERENCE: 314572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Atlantic cod
US-09-938-269-16

Query Match      56.1%; Score 60; DB 9; Length 25;
Best Local Similarity 50.0%; Pred. No. 0.0073;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1  IVGGYTCGANTVPQVSLNS 20
      ||||| : : : : :
      ||||| : : : : :

DB      1  IVGGYECTRHQAQVSLNS 20

RESULT 13
US-10-021-368-10
; Sequence 10, Application US/10021368
; Patent No. US20020106367A1
; GENERAL INFORMATION:
; APPLICANT: Band, Viola
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; SOFTWARE: Patent Release/MS-DOS
; CURRENT APPLICATION DATA: US/10/021,368
; APPLICATION NUMBER: US/10/021,368
; FILING DATE: 12-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/201,038
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELETYPE: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids

```

Wed Feb 12 11:59:33 2003

Query Match 45.8%; Score 49; DB 10; Length 235;
Best Local Similarity 45.0%; Pred. NO. 3.1;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGYTCGANTVPYQVSLNS 20
Db 87 IVGGWECERKISQPMQVAVYS 106

Search completed: February 12, 2003, 10:31:04
Job time : 4.04478 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:03:55 ; Search time 8.7134 Seconds
(without alignments)
198.395 Million cell updates/sec

Title: US-10-036-371-3

Sequence: 1 CXFAPXPKXTGY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSZ/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	86.4	13	22 AAB31575	Amino terminal seq
2	40	67.8	13	20 AAY50209	Neutrophil-activat
3	40	67.8	13	22 AAB31577	Amino terminal seq
4	40	67.8	14	21 AAB24191	Neutrophil-activat
5	40	67.8	24	21 AAY95996	Bovine chymotrypsi
6	40	67.8	24	21 AAY95996	Amino acid sequenc
7	40	67.8	26	23 AAU82718	Neutrophil-activat
8	39	66.1	15	20 AAY50208	Neutrophil-activat
9	39	66.1	15	20 AAY50210	Neutrophil-activat
10	39	66.1	811	12 AAR12944	Plasminogen mutcin

11	39	66.1	813	12 AAR12936	Plasminogen mutcin
12	39	66.1	813	12 AAR12937	Plasminogen mutcin
13	38	64.4	626	22 ABB61150	Drosophila melanog
14	36	61.0	51	22 AAU54358	Propionibacterium
15	36	61.0	106	23 ABP09202	Human ORFX protein
16	36	61.0	264	21 AAB11711	Mouse serine prote
17	35	59.3	13	20 AAY50212	Neutrophil-activat
18	35	59.3	44	21 AAB64742	Human 5' EST relat
19	35	59.3	55	23 ABP03011	Human ORFX protein
20	35	59.3	188	22 AAU61671	Propionibacterium
21	35	59.3	192	21 AAB54077	Human pancreatic c
22	35	59.3	264	21 AAB11710	Human serine prote
23	35	59.3	87	23 AAY01713	Human chordin (CH
24	35	59.3	885	23 AAY01713	A human huchordin
25	35	59.3	885	23 AAY01956	Human PRO243 Hom
26	35	59.3	936	22 AAE03735	Murine mature chor
27	35	59.3	948	22 AAE03736	Murine chordin (CH
28	35	59.3	948	23 AAU75168	Mouse chordin poly
29	35	59.3	954	19 AAY48978	Mature human chori
30	35	59.3	954	20 AAV78211	Human PRO243 prote
31	35	59.3	954	21 AAB01312	Human PRO243 poly
32	35	59.3	954	23 AAU86132	Human PRO243 poly
33	35	59.3	955	22 AAE12889	Human chordin prot
34	35	59.3	962	22 AAE03734	Murine chordin (CH
35	34	57.6	110	21 AAB40871	Human ORFX ORF635
36	34	57.6	110	23 ABP32707	Human ORF1680 prot
37	34	57.6	112	21 AAG60981	Arabidopsis thalia
38	34	57.6	117	21 AAG60978	Arabidopsis thalia
39	34	57.6	117	21 AAG60980	Arabidopsis thalia
40	34	57.6	120	22 AAB56845	Arabidopsis thalia
41	34	57.6	121	21 AAG60977	Arabidopsis thalia
42	34	57.6	121	21 AAG60977	Arabidopsis thalia
43	34	57.6	121	21 AAG60977	Arabidopsis thalia
44	34	57.6	183	21 AAG04375	Arabidopsis thalia
45	34	57.6	183	21 AAG60976	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ID AAB31575 standard; peptide; 13 AA.

XX AAB31575;

XX 20-APR-2001 (first entry)

XX Amino terminal sequence of cod trypsin A chain.

XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;
XX arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;
XX rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
XX systemic lupus erythematosus; phlebitis; tendinitis; rash; psoriasis;
XX acne; eczema; facial seborrheic eczema; foreskin infection;
XX athlete's foot; fistulae infection; ulcer; navel infection;
XX scar; keloid; boil; wart; allergic itch; hemorroid; wound;
XX fungal infection; autoimmune disease.

XX Gadus callarias.

XX Key Location/Qualifiers

FT Misc-difference 3 /label= Arg, Ser

FT Misc-difference 9 /label= Ser, Gln

FT Misc-difference 10 /label= Val, Gln

FT Misc-difference 10 /label= Ile, Val

XX MC200078332-A2.

PD 28-DEC-2000.
 XX
 XX
 XX 15-JUN-2000; 2000MO-1S00005.
 XX
 XX 18-JUN-1999; 99IS-0005086.
 XX
 XX (BJAR/) BJARNASON J B.
 XX
 XX Bjarnason JB;
 XX
 XX WPI; 2001-091493/10.
 XX
 XX Fish serine proteinase, useful as a cosmetic, medicament for treating
 PT eczema, psoriasis, arthritis, and in the manufacture of the medicament
 PT for treating, preventing pathogenic diseases involving receptor
 PT mediated binding -
 XX
 XX Disclosure; Page 5; 38pp; English.
 XX
 XX The specification describes a fish serine proteinase. The proteinases
 CC are useful as medicaments, for treating and preventing a disease in a
 CC human or an animal such as pain, acute inflammation, chronic
 CC inflammation, arthritis, inflamed joints, bursitis, osteoarthritis,
 CC rheumatoid arthritis, juvenile rheumatoid arthritis, septic arthritis,
 CC fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, rash,
 CC psoriasis, acne, eczema, facial seborrheic eczema, eczema of the hands,
 CC face or neck, foreskin infections, athlete's foot, fistulae infections,
 CC keloids, boils, warts and allergic itch, hemorrhoids, wrinkles, scars,
 CC infections, wounds from burns, fungal infections and immunological and
 CC autoimmune diseases. They are also useful for removing dead or peeling
 CC skin from cheeks, healthy skin, and for treating or preventing a
 CC disease in which pathogenesis is caused by bacteria, virus, fungus,
 CC parasite or a protozoan. A receptor mediated binding is involved.
 CC The present sequence represents the amino terminal of cod trypsin A
 CC chain.
 XX
 XX Sequence 13 AA;
 XX
 XX Query Match 86.4%; Score 51; DB 22; Length 13;
 XX Best Local Similarity 100.0%; Pred. No. 0.011;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 CGXPAIXPXXTGY 13
 XX |||||
 XX DB 1 CGXPAIXPXXTGY 13
 XX
 XX RESULT 2
 XX ID AAY50209 standard; Peptide; 13 AA.
 XX AC AAY50209;
 XX
 XX DT 12-JAN-2000 (first entry)
 XX
 XX DE Neutrophil-activating pancreatic derived peptide 9.
 XX
 XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
 KW inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
 KW organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
 KW hemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
 KW trauma; protease inhibitor; hypertension; sepsis.
 XX
 XX Bos taurus.
 XX
 XX WO9946367-A2.
 XX
 XX 16-SEP-1999.
 XX
 XX 11-MAR-1999; 99MO-US05247.
 XX
 XX 11-MAR-1998; 98US-0038894.

(CELL-) CELL ACTIVATION INC.
 (REGC) UNIV CALIFORNIA.
 (SCRI) SCRIPPS RES INST.
 Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kierler E;
 WPI, 1999-580234/49.
 Use of cell activating compositions in developing products for
 diagnosis and treatment of e.g. cardiovascular, inflammatory,
 autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection,
 diabetes, stroke or ischemia -
 PS Example 9; Page 182; 184pp; English.
 XX
 XX This invention describes a novel method for the use and preparation of
 CC cell activating compositions which involves preparing a cell activating
 CC composition comprising (a) homogenizing pancreatic tissue in buffer at
 CC about neutral or higher pH to produce a homogenate; (b) removing
 CC particulates from the homogenate; (c) optionally incubating the
 CC resulting homogenate, with particulates removed, with a protease; and
 CC (d) fractionating the homogenate and selecting fractions that exhibit
 CC cell activation activity. The methods can be used for improving
 CC treatment outcome or reducing risk of treatment of e.g. cardiovascular
 CC disease, inflammatory disease, trauma, autoimmune diseases, arthritis,
 CC organ rejection, diabetes and diabetic complications, stroke, ischemia,
 CC Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic
 CC retinopathy, diabetes, venous insufficiency, unstable angina or trauma.
 CC They can be used in the veterinary treatment of a non-human subject.
 CC Protease inhibitors can be used to lower cell activation resulting from
 CC these diseases and deficiencies. The detection of an elevated level of
 CC hydrogen peroxide can be used to detect an inflammatory condition. An
 CC elevated level of hydrogen peroxide in plasma or whole blood and in the
 CC presence of superoxide dismutase (SOD) indicates leukocyte up
 CC regulation, e.g. indicative of the onset of an acute cardiovascular
 CC disorders, such as disease onset or ischemic complications. An elevated
 CC level of hydrogen peroxide in plasma or whole blood and a low level in
 CC the presence of SOD is indicative of a chronic or immune compromised
 CC condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
 CC used in the method of the invention.
 XX
 XX Sequence 13 AA;
 XX
 XX Query Match 67.8%; Score 40; DB 20; Length 13;
 XX Best Local Similarity 58.3%; Pred. No. 1;
 XX Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 XX
 XX QY 1 CGXPAIXPXXTG 12
 XX |||||
 XX DB 1 CGVPAIQPVLSG 12
 XX
 XX RESULT 3
 XX AAB31577
 XX ID AAB31577 standard; peptide; 13 AA.
 XX AC AAB31577;
 XX
 XX XX 20-APR-2001 (first entry)
 XX
 XX DE Amino terminal sequence of bovine trypsin A chain.
 XX
 XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;
 KW arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;
 KW rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
 KW systemic lupus erythematosus; phlebitis; tendinitis; rash; psoriasis;
 KW acne; eczema; facial seborrheic eczema; foreskin infection;
 KW athlete's foot; fistulae infection; ulcer; navel infection; wrinkle;
 KW scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;
 KW fungal infection; autoimmune disease.
 XX
 XX Bos sp.

XX WO200078332-A2.
 XX PD 28-DEC-2000.
 XX PF 15-JUN-2000; 2000WO-1500005.
 XX PA (SCRI) UNIV CALIFORNIA.
 XX PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
 XX WPI; 1999-580234/49.
 XX Use of cell activating compositions in developing products for
 XX diagnosis and treatment of ligamentous, cardiovascular, inflammatory,
 XX autoimmune, Alzheimer's disease, trauma, arthritis, organ rejection,
 XX diabetes, stroke or ischemia.
 XX Example 9, Page 182; 184pp; English.
 XX This invention describes a novel method for the use and preparation of
 XX cell activating compositions which involves preparing a cell activating
 XX composition comprising (a) homogenizing pancreatic tissue in buffer at
 XX about neutral or higher pH to produce a homogenate; (b) removing the
 XX particulates from the homogenate; (c) optionally incubating the
 XX resulting homogenate, with particulates removed, with a protease; and
 XX (d) fractionating the homogenate and selecting fractions that exhibit
 XX cell activation activity. The methods can be used for improving
 XX treatment outcome or reducing risk of treatment of e.g. Cardiovascular
 XX disease, inflammatory disease, trauma, autoimmune diseases, arthritis,
 XX organ rejection, diabetes and diabetic complications, stroke, ischemic
 XX Alzheimer's disease, myocardial infarction, metabolic syndrome, retinopathy,
 XX diabetes, venous insufficiency, a non-human subject from
 XX protease inhibitors can be used to lower cell activation resulting from
 XX these diseases and deficiencies. The detection of an elevated level of
 XX hydrogen peroxide can be used to detect an inflammatory condition. An
 XX elevated level of hydrogen peroxide in plasma or whole blood and in the
 XX presence of superoxide dismutase (SOD) indicates leukocyte up
 XX regulation, e.g. indicative of the onset of an acute cardiovascular
 XX disorders, such as disease onset or ischemic complications. An elevated
 XX level of hydrogen peroxide in plasma or whole blood and a low level in
 XX the presence of SOD is indicative of a chronic or immune compromised
 XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
 XX used in the method of the invention.

XX SQ Sequence 13 AA;
 Query Match 67.8%; Score 40; DB 22; Length 13;
 Best Local Similarity 58.3%; Pred. No. 1;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CXGPAIXPXXTG 12
 DB 1 CGVPAIQPVLSG 12
 |||||
 |||||

RESULT 4
 AAY50207
 ID AAY50207 standard; Peptide, 15 AA.
 AC AAY50207;
 XX 12-JAN-2000 (first entry)
 XX Neutrophil-activating pancreatic derived peptide 7.
 XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
 XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
 XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
 XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
 XX trauma; protease inhibitor; hypertension; sepsis.
 XX Bos taurus.
 XX WO9946367-A2.
 XX 16-SEP-1999.

PF 11-MAR-1999; 99WO-US05247.
 XX PD 11-MAR-1998; 98US-0038894.
 XX PA (CELL-) CELL ACTIVATION INC.
 XX (REGC) UNIV CALIFORNIA.
 XX PA (SCRI) SCRIPPS RES INST.
 XX PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
 XX WPI; 1999-580234/49.
 XX Use of cell activating compositions in developing products for
 XX diagnosis and treatment of ligamentous, cardiovascular, inflammatory,
 XX autoimmune, Alzheimer's disease, trauma, arthritis, organ rejection,
 XX diabetes, stroke or ischemia.
 XX Example 9, Page 182; 184pp; English.
 XX This invention describes a novel method for the use and preparation of
 XX cell activating compositions which involves preparing a cell activating
 XX composition comprising (a) homogenizing pancreatic tissue in buffer at
 XX about neutral or higher pH to produce a homogenate; (b) removing the
 XX particulates from the homogenate; (c) optionally incubating the
 XX resulting homogenate, with particulates removed, with a protease; and
 XX (d) fractionating the homogenate and selecting fractions that exhibit
 XX cell activation activity. The methods can be used for improving
 XX treatment outcome or reducing risk of treatment of e.g. Cardiovascular
 XX disease, inflammatory disease, trauma, autoimmune diseases, arthritis,
 XX organ rejection, diabetes and diabetic complications, stroke, ischemic
 XX Alzheimer's disease, myocardial infarction, metabolic syndrome, retinopathy,
 XX diabetes, venous insufficiency, a non-human subject from
 XX protease inhibitors can be used to lower cell activation resulting from
 XX these diseases and deficiencies. The detection of an elevated level of
 XX hydrogen peroxide can be used to detect an inflammatory condition. An
 XX elevated level of hydrogen peroxide in plasma or whole blood and in the
 XX presence of superoxide dismutase (SOD) indicates leukocyte up
 XX regulation, e.g. indicative of the onset of an acute cardiovascular
 XX disorders, such as disease onset or ischemic complications. An elevated
 XX level of hydrogen peroxide in plasma or whole blood and a low level in
 XX the presence of SOD is indicative of a chronic or immune compromised
 XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
 XX used in the method of the invention.

XX SQ Sequence 15 AA;
 Query Match 67.8%; Score 40; DB 20; Length 15;
 Best Local Similarity 58.3%; Pred. No. 1,2;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CXGPAIXPXXTG 12
 DB 1 CGVPAIQPVLSG 12
 |||||
 |||||

RESULT 5
 AAB54191
 ID AAB54191 standard; Protein; 146 AA.
 XX AAB54191;
 XX 09-MAR-2001 (first entry)
 XX Human pancreatic cancer antigen protein sequence SEQ ID NO:643.
 XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 XX detection; diagnosis; identification; cytostatic; neuroprotective;
 XX neoplastic; immunomodulatory; relaxant; contraceptive; synaectological;
 XX antiinflammatory; cardiant; gene therapy; chromosome mapping;
 XX linkage analysis; tissue identification; tissue typing; forensic;
 XX neural; immune system; muscular; reproductive; gastrointestinal;
 XX pulmonary; cardiovascular; renal; proliferative.

XX OS Homo sapiens.
 XX PN WO200055320-A1.
 XX PD 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US05989.
 XX PR 12-MAR-1999; 99US-0124270.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Ruben SM;
 XX DR WPI; 2000-579444/54.
 XX DR N-PSDB; AAC98596.
 XX PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 XX PT treating, or ameliorating a medical condition, particular pancreatic
 XX PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 XX Claim 11; Page 1081; 1379pp; English.
 XX CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
 XX CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 XX CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 XX CC neuroprotective, chemoprotective, immunomodulatory, relaxant, contraceptive,
 XX CC gynecological, and anti-inflammatory activities, and can be used
 XX CC in gene therapy. The polypeptides and proteins can be used for
 XX CC preventing, treating or ameliorating a medical condition or in assays
 XX CC for diagnosing a pathological condition. A nucleic acid sequence of one in a
 XX CC subject. Binding partners to the proteins and the polypeptides can be used to
 XX CC detect, treat or prevent pancreatic disorders, especially cancer.
 XX CC Agonists and antagonists to the antigens can be screened for. The
 XX CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 XX CC acid hybridization probes that can be used in chromosome mapping, linkage
 XX CC analysis, tissue identification and/or typing and a variety of forensic
 XX CC and diagnostic methods. The proteins can be used to generate antibodies
 XX CC which are used to purify, detect and target the polypeptides, including
 XX CC both in vivo and in vitro diagnostic and therapeutic methods. The
 XX CC proteins can be used to treat or prevent neural, immune system, muscular,
 XX CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 XX CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 XX CC sequences used in the exemplification of the present invention.
 XX SQ Sequence 146 AA;
 Query Match 67.8%; Score 40; DB 21; Length 146;
 Best Local Similarity 58.3%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGVPAIXPXXTG 12
 DB 23 CGVPAHPVLVG 34
 RESULT 6
 AA199596
 ID AA199596 standard; protein; 245 AA.
 XX AA199596;
 XX AA199596;
 XX DT 13-SEP-2000 (first entry)
 XX DE Bovine chymotrypsinogen A.
 XX KW Bovine; plasminogen activator; cardiant; thrombolytic;
 XX KW heart attack; stroke; blood clotting disorder.
 XX XX Bos taurus.
 XX OS . . .

PN WO200032759-A1.
 XX 08-JUN-2000.
 XX PF 06-MAY-1999; 99WO-US09991.
 XX PR 02-DEC-1998; 98US-0110588.
 XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 XX PI Lin X, Zhang XC, Tang JUN;
 XX DR WPI; 2000-422975/36.
 XX PT Polypeptide with plasminogen activator activity useful as thrombolytic
 XX PT agent for treating blood clot disorders e.g. heart attack, comprises 10
 XX PT amino acid peptide fragment for recognition or activation of
 XX PT plasminogen -
 XX PS
 XX Disclosure; Page 35-36; 41pp; English.
 XX CC The present sequence is bovine chymotrypsinogen. It was included in
 XX CC a review of sequence homologies of several plasminogen activators.
 XX CC Plasminogen is the principal serine protease zymogen in the
 XX CC extracellular fluids of vertebrates. Its active form, plasmin, is
 XX CC implicated in pericellular proteolysis associated with a wide range of
 XX CC physiological and pathological processes. Plasminogen expression is
 XX CC regulated by plasminogen activators which hydrolyse a peptide bond in
 XX CC plasminogen to convert it to plasmin or form tight binding complexes
 XX CC with plasminogen to spontaneously convert it to plasmin. The sequence
 XX CC homology analysis has identified a six amino acid peptide involved in
 XX CC plasminogen activation. This peptide is particularly useful when inserted
 XX CC between amino acid residues 644 and 645 of full length human
 XX CC plasminogen. Novel plasminogen activators have been made based upon the
 XX CC plasminogen activation/recognition site of plasminogen binding proteins.
 XX CC The polypeptides are useful in preparing thrombolytic agents for treating
 XX CC blood clotting disorders such as heart attack.
 XX SQ Sequence 245 AA;
 Query Match 67.8%; Score 40; DB 21; Length 245;
 Best Local Similarity 58.3%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGVPAIXPXXTG 12
 DB 1 CGVPAHPVLVG 12
 RESULT 7
 AAU82738
 ID AAU82738 standard; Protein; 263 AA.
 XX AAU82738;
 XX AAU82738;
 XX DT 23-APR-2002 (first entry)
 XX DE Amino acid sequence of novel human protease #37.
 XX KW Human; protease; cancer; immune-related disorder; cardiovascular disease;
 XX KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
 XX KW nervous system disorder; sexual dysfunction; pain; mood disorder;
 XX KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
 XX KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
 XX KW ocular disease; cytostatic; enzyme.
 XX OS Homo sapiens.
 XX PN WO200200860-A2.
 XX PD 03-JAN-2002.
 XX PT 26-JUN-2001; 2001WO-US20171.

XX 11-MAR-1999; 99WO-US05247.
 XX 11-MAR-1998; 98US-0038894.
 XX (CELL-) CELL ACTIVATION INC.
 PA (REGC) UNIV CALIFORNIA.
 PA (SCRI) SCRIPPS RES INST.
 XX
 FI Soughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
 DR WPI; 1999-580234/49.
 XX
 XX Use of cell activating compositions in developing products for
 PT diagnosis and treatment of e.g. cardiovascular, inflammatory,
 PT autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection,
 PT diabetes, stroke or ischemia -
 XX
 PS Example 9; Page 182; 184pp; English.
 XX
 XX This invention describes a novel method for the use and preparation of
 CC cell activating compositions which involves preparing a cell activating
 CC composition comprising (a) homogenizing pancreatic tissue in buffer at
 CC physiological or higher pH to produce a homogenate; (b) removing
 CC particulates from the homogenate; (c) optionally incubating the
 CC resulting homogenate with particulates removed; with a protease; and
 CC (d) fractionating the homogenate and selecting fractions that exhibit
 CC cell activation activity. The method can be used for improving
 CC treatment outcome or reducing risk of treatment of e.g. cardiovascular
 CC disease, inflammatory disease, trauma, autoimmune disease, arthritis,
 CC organ rejection, diabetes and diabetic complications, stroke, ischemia,
 CC Alzheimer's disease, myocardial infarction, hemorrhagic shock, diabetic
 CC retinopathy, diabetes, venous insufficiency, unstable angina or claudication.
 CC They can be used in the veterinary treatment of a non-human subject.
 CC Protease inhibitors can be used to lower cell activation resulting from
 CC these diseases and deficiencies. The detection of an elevated level of
 CC hydrogen peroxide can be used to detect an inflammatory condition. An
 CC elevated level of hydrogen peroxide in plasma or whole blood and in the
 CC presence of superoxide dismutase (SOD) indicates leukocyte up
 CC regulation, e.g. indicative of the onset of an acute cardiovascular
 CC disorders, such as disease onset or ischemic complications. An elevated
 CC level of hydrogen peroxide in plasma or whole blood and a low level in
 CC the presence of SOD is indicative of a chronic or immune compromised
 CC condition e.g. hypertension or sepsis. AA450201-Y50334 represent peptides
 CC used in the method of the invention.
 XX
 XX Sequence 15 AA;
 SQ
 Query Watch 66.1%; Score 39; DB 20; Length 15;
 Best Local Similarity 58.3%; Pred. No. 1.8;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGXPALXPXXGT 12
 |||||
 DB 1 CGVPAIPVLSG 12

RESULT 10
 AAR12944
 ID AAR12944 standard; Protein; 811 AA.
 AC AAR12944;
 XX
 DT 23-SEP-1991 (first entry)
 XX
 DE Plasminogen mutcin T14 with thrombin cleavage site.
 XX
 KW protease; fibrinolysis; blood clotting; thrombosis.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site /note= "recognised by thrombin"
 XX
 XX . ,

PN WO9109118-A.
 XX
 PD 27-JUN-1991.
 XX
 PF 07-DEC-1990; 90WO-GB01912.
 XX
 PR 07-DEC-1989; 89GB-0027722.
 XX
 PR 07-DEC-1990; 90WO-GB01911.
 XX
 PA (BRBI-) BRIT BIO-TECHN LTD.
 XX
 XX Dawson KM, Edwards RM, Forman JM;
 PI WPI; 1991-208145/28.
 XX
 DR N-PSDB; AAQ12553.
 XX
 DR Activatable fibrinolytic and antithrombotic proteins - activated by
 PT e.g. factor Xa, thrombin or activated protein C
 XX
 PS Claim 9; Fig 2 and Fig 5; 73pp; English.
 XX
 XX This protein is a plasminogen mutant cleavable by thrombin.
 CC Activation is localised to the thrombus because cleavage to plasmin
 CC is by an enzyme of the blood clotting pathway. Compositions
 CC comprising the mutant plasminogen are used for treatment or
 CC prevention of thrombosis, etc.
 CC See AAQ12542-Q12558.
 XX
 SQ Sequence 811 AA;
 Query Match 66.1%; Score 39; DB 12; Length 811;
 Best Local Similarity 46.2%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CGXPALXPXXGT 13
 |||||
 DB 567 CGKQVEPKCDY 579
 RESULT 11
 AAR12936
 ID AAR12936 standard; Protein; 813 AA.
 AC AAR12936;
 XX
 DT 23-SEP-1991 (first entry)
 XX
 DE Plasminogen mutcin X5 with factor Xa cleavage site.
 XX
 KW protease; fibrinolysis; blood clotting; thrombosis.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site /label= Factor Xa cleavage site
 XX
 PN WO9109118-A.
 XX
 PD 27-JUN-1991.
 XX
 PF 07-DEC-1990; 90WO-G001912.
 XX
 PR 07-DEC-1989; 89GB-0027722.
 XX
 PR 07-DEC-1990; 90WO-GB01911.
 XX
 PA (BRBI-) BRIT BIO-TECHN LTD.
 XX
 XX Dawson KM, Edwards RM, Forman JM;
 PI WPI; 1991-208145/28.
 XX
 DR N-PSDB; AAQ12545.
 XX
 DR Activatable fibrinolytic and antithrombotic proteins - activated by
 PT e.g. factor Xa, thrombin or activated protein C
 XX

```

XX PS Claim 7; Fig 2 and Fig 4; 73pp; English.
XX CC This protein is a plasminogen mutant cleavable by Factor Xa.
XX CC As by activation of the mutant thrombus because cleavage to plasmin
XX CC is by an enzyme of the blood clotting pathway. Compositions
XX CC comprising the mutant plasminogen are used for treatment or
XX CC prevention of thrombosis, etc.
XX CC See AAQ12542-Q12558.
XX CC
XX SQ Sequence 813 AA;
XX
XX Query Match 66.1%; Score 39; DB 12; Length 813;
XX Best Local Similarity 46.2%; Pred. No. 92;
XX Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 CGXPAIXPXTGY 13
XX |||:|:|
XX Db 567 CGKQVEPKCGY 579
XX
XX RESULT 12
XX AARI12917
XX ID AARI12937 standard; Protein; 813 AA.
XX AC AARI12937;
XX DT 23-SEP-1991 (first entry)
XX DE Plasminogen mutcin X6 with factor Xa cleavage site.
XX KW protease; fibrinolysis; blood clotting; thrombosis.
XX FX Key Location/Qualifiers
XX FT Cleavage-site 716..585
XX FN /label= Factor Xa cleavage site
XX PN WO9109118-A.
XX PD 27-JUN-1991.
XX PF 07-DEC-1990; 90WO-GO01912.
XX PR 07-DEC-1989; 89GB-0027722.
XX PR 07-DEC-1990; 90WO-GS01911.
XX PA (BRRI-) BRIT BIO-TECHN LTD.
XX PI Dawson KM, Edwards RM, Forman JM;
XX DR WPI; 1991-208145/28.
XX DR N-PSDB; AAQ12546.
XX YX Activatable fibrinolytic and antithrombic proteins - activated by
XX PT e.g. factor Xa, thrombin or activated protein C
XX PS Claim 7; Fig 2 and fig 4; 73pp; English.
XX
XX CC This protein is a plasminogen mutant cleavable by Factor Xa.
XX CC Activation is localised to the thrombus because cleavage to plasmin
XX CC is by an enzyme of the blood clotting pathway. Compositions
XX CC comprising the mutant plasminogen are used for treatment or
XX CC prevention of thrombosis, etc.
XX CC See AAQ12542-Q12558.
XX CC
XX SQ Sequence 813 AA;
XX
XX Query Match 66.1%; Score 39; DB 12; Length 813;
XX Best Local Similarity 46.2%; Pred. No. 92;
XX Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 CGXPAIXPXTGY 13
XX |||:|:|
XX Db 567 CGKQVEPKCGY 579
XX
XX RESULT 13
XX ABB61150
XX ID ABB61150 standard; Protein; 626 AA.
XX AC ABB61150;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 10242.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WC200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABU92553.
XX YX New isolated nucleic acid detection reagent for detecting 1003 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 10242; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABU16176-ABU16311), expressed DNA
XX CC sequences (ABU01840-ABU16175) and the encoded proteins
XX CC (AB85793-AB85795).
XX CC The sequences disclosed for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 626 AA;
XX
XX Query Match 64.4%; Score 38; DB 22; Length 626;
XX Best Local Similarity 41.7%; Pred. No. 1.1e+02;
XX Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 CGXPAIXPXTGY 12
XX |||:|:|
XX Db 406 CGTSPSLPYSSG 417
XX
XX RESULT 14
XX AAU54358
XX ID AAU54358 standard; Protein; 51 AA.
XX AC AAU54358;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #15254.
XX

```

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 KW WO200181581-A2.
 PN
 XX
 PD 01-NOV-2001.
 XX
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Skelly YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX
 DR WPI; 2001-616774/71.
 XX
 XX N-PSDB; AB859564.
 XX
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX
 PS Example 1, SEQ ID No 15553; 1069pp; English.
 XX
 CC Sequences AAU39105-BAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome, synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis of bone, and endophthalmitis.
 CC P. acnes is also involved in infections of bone, and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting the
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 SQ Sequence 51 AA;
 Query Match 61.0%; Score 36; DB 22; Length 51;
 Best Local Similarity 54.5%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGXPAIXPXYT 11
 DB 35 CQNFVAPVST 45
 RESULT 15
 ID ABP09202 standard; Protein; 106 AA.
 XX
 AC ABP09202;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Human ORFX protein sequence SEQ ID NO:18386.
 XX

KW Human, open reading frame; ORFX, gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW hepatocellular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertonic hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2001192523-A2.
 PN
 XX
 XX 06-DEC-2001.
 PD
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX
 XX 30-MAY-2000; 2000US-206132P.
 PR
 PR 29-AUG-2000; 2000US-228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 XX Shmkets RA, Leach MD;
 PI
 XX WPI; 2002-106308/14.
 DR
 DR N-PSDB; ABN24954.
 XX
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 XX Disclosure; SEQ ID 18386; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN15765 encode the human ORFX
 CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, myasthenia, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, myasthenia, graft-versus-host
 CC storage disease, various immune deficiencies and disorders, rheumatoid
 CC disease, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 SQ Sequence 106 AA;
 Query Match 61.0%; Score 36; DB 23; Length 106;
 Best Local Similarity 54.5%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 GXPAIXPXYTG 12
 DB 7 GKPAIXPAHTG 17
 Search completed: February 12, 2003, 10:22:23
 Job time : 9.73134 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 3.2597 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-3

Perfect score: 59

Sequence: 1 CGSPAIPXXTGY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	52	88.1	244	2	S72219	chymotrypsin B - A
2	52	88.1	263	2	S47537	chymotrypsin (EC 3
3	43	72.9	1365	2	T30198	alkaline phosphata
4	41	69.5	28	2	A61539	chymotrypsin (EC 3
5	40	67.8	245	1	KY808	chymotrypsin (EC 3
6	40	67.8	245	1	KY808	chymotrypsin (EC 3
7	40	67.8	245	1	KY808	chymotrypsin (EC 3
8	40	67.8	245	1	KY808	chymotrypsin (EC 3
9	40	67.8	245	1	KY808	chymotrypsin (EC 3
10	39	66.1	16	2	C63414	chymotrypsin (EC 3
11	39	66.1	17	2	B61414	chymotrypsin (EC 3
12	37	62.7	542	2	A82965	hypothetical prote
13	36	61.0	321	2	T07043	probable epoxide h
14	36	61.0	350	2	AD3170	two component sens
15	36	61.0	373	2	AH3382	transcription repr
16	35	59.3	111	2	C71227	hypothetical prote
17	35	59.3	120	2	A71017	hypothetical prote
18	35	59.3	126	2	A23473	chymotrypsin-like
19	35	59.3	264	2	I38136	chymotrypsin-like
20	35	59.3	887	2	AG0535	ClpB-like protein
21	35	59.3	4077	2	T17484	hypothetical prote
22	34	57.6	117	2	G75115	hypothetical prote
23	34	57.6	118	2	A24777	homeotic protein H
24	34	57.6	157	2	F84586	hypothetical prote
25	34	57.6	157	2	F84586	conserved hypotet
26	34	57.6	265	2	S80260	serine proteinase
27	34	57.6	265	2	S80260	serine proteinase
28	34	57.6	282	2	A84341	hypothetical prote
29	34	57.6	312	1	Q08E07	US10 protein - hum

30 34 57.6 378 2 A33271 manganese peroxida
31 34 57.6 442 2 H71410 probable RNA helic
32 34 57.6 479 2 H90042 hypothetical prote
33 34 57.6 558 2 I56545 glycan precursor
34 34 57.6 571 2 T52576 ATP-dependent RNA
35 34 57.6 643 2 C90035 hypothetical prote
36 33 55.9 107 2 B72498 hypothetical prote
37 33 55.9 213 2 C89865 conserved hypotet
38 33 55.9 215 2 T22572 hypothetical prote
39 33 55.9 231 2 S37108 cuticlin 2 - Caeno
40 33 55.9 245 2 S06329 chlorophyll a/b-bi
41 33 55.9 246 2 S00443 chlorophyll a/b-bi
42 33 55.9 333 2 AD3284 proline racemase (

ALIGNMENTS

RESULT 1

S72219 chymotrypsin B - Atlantic cod (fragments)

C:Species: Gadus morhua (Atlantic cod)

C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Aug-1998

C:Accession: S72219

R:Leth-Larsen, R.; Ageirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.

Biochim. Biophys. Acta 1297, 49-56, 1996

A>Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.

A:Reference number: S72219; MUID:96439045; PMID:8841380

A:Accession: S72219

A>Status: preliminary

A:Residue type: 154, 150, 244, <LET>

A:Protein family: trypsin; trypsin homology

C:Superfamily: trypsin; trypsin homology

F:15-237/Domain: trypsin homology <TRY>

Query Match 88.1%; Score 52; DB 2; Length 244;

Best Local Similarity 69.2%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGSPAIPXXTGY 13

DB 1 CGSPAIPQVTGY 13

RESULT 2

S47537 chymotrypsin (EC 3.4.21.1), precursor - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C>Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999

C:Accession: S47537; S41163

R:Gundundersdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.

Biochim. Biophys. Acta 1219, 211-214, 1994

A>Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.

A:Reference number: S47537; MUID:94168860; PMID:8086467

A:Accession: S47537

A:Molecule type: mRNA

A:Residues: 1-263 <GUD>

C:Cross-references: EMBL:X78490; NID:9468750; PIDN:CAA55242.1; PID:9468751

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-263/Product: chymotrypsin #status predicted <MAT>

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 88.1%; Score 52; DB 2; Length 263;

Best Local Similarity 69.2%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGSPAIPXXTGY 13

```

Db      19  CGRPAISPVITG 31
|||||
RESULT 3
T30198
C:Species: Volvox carteri
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30198
R:Halimann, A.
J. Biol. Chem. 274, 1691-1697, 1999
A:Title: Enzymes in the extracellular matrix of Volvox: an inducible, calcium-dependent
A:Reference number: #20772; MUID:99098915; PMID:9880549
A:Accession: T30198
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1365 <HAU>
A:Cross-references: EMBL:AJ012458; NID:94160585; PID:CAA10030.1; PID:94160586
A:Experimental source: ssp. nagariensis, strain HK 10
C:Genetic:
A:Gene: phox
A:Function:
C:Description: hydrolyzes phosphate esters
C:Superfamily: Volvox carteri alkaline phosphatase phox
C:Keywords: magnesium; metalloprotein; phosphoric monoester hydrolase; zinc
Query Match      72.9%; Score 43; DB 2; Length 1365;
Best Local Similarity 53.8%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy      1  CGXPALXPXTGY 13
|||||
Db      1063  CGTATPAAGY 1075
|||||

RESULT 4
A61529
chymotrypsin (EC 3.4.21.1) - Atlantic cod (fragments)
C:Species: Gadus morhua (Atlantic cod)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 11-May-2000
C:Accession: A61529
R:Aggelarson, B.; Bjarnason, J.B.
Comp. Biochem. Physiol. B 99, 327-335, 1991
A:Title: Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua)
A:Reference number: A61529; MUID:92111252; PMID:1764912
A:Accession: A61529
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-28 <ASG>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
Query Match      69.5%; Score 41; DB 2; Length 28;
Best Local Similarity 58.3%; Pred. No. 0.19;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy      1  CGXPALXPXTG 12
|||||
Db      1  CGSPAQPVISS 12
|||||

RESULT 5
KYBOA
chymotrypsin (EC 3.4.21.1) B precursor - bovine
N:Alternate names: chymotrypsinogen B
C:Species: Bos primigenius taurus (cattle)
C:Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 18-Jul-1997
C:Accession: A00953
R:Smillie, L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.
Nature 218, 343-346, 1968
A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen
A:Reference number: A00953; MUID:60238908; PMID:5649671

```

```

A:Accession: A00953
A:Molecule type: protein
A:Residues: 1-245 <SMI>
C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acinar cells of the pancreas. The first activation cleavage, leading to pi-chymotrypsin B, occurs in the duodenum.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: propeptide #status experimental <PRO>
F:16-245/Product: chymotrypsin B #status experimental <PRO>
F:16-238/Domain: trypsin homology <TRY>
F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental
F:57,102,195/Active site: His, Asp, Ser #status experimental
Query Match      67.8%; Score 40; DB 1; Length 245;
Best Local Similarity 58.3%; Pred. No. 2.4;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy      1  CGXPALXPXTG 12
|||||
Db      1  CGSPAQPVISS 12
|||||

RESULT 6
KYBOA
chymotrypsin (EC 3.4.21.1) A precursor - bovine
N:Alternate names: chymotrypsinogen A
C:Species: Bos primigenius taurus (cattle)
C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999
C:Accession: A90235; A93158; S29650; A00952
R:Brown, J.R.; Hartley, B.S.
Biochem. J. 101, 214-228, 1966
A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide
A:Reference number: A90235; MUID:67181721; PMID:5971783
A:Accession: A90235
A:Molecule type: protein
A:Residues: 1-101, N', 103-245 <BRO>
R:Blow, D.M.; Birktott, J.J.; Hartley, B.S.
Nature 221, 336-340, 1969
A:Title: Role of a buried acid group in the mechanism of action of chymotrypsin.
A:Reference number: A93158; MUID:69106467; PMID:5764336
A:Accession: A93158
C:Keywords: annotation; trypsin; residue 102
R:McLoun, B.; Klum, T.; Kesteven, V.; Moser, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Sorn,
Biochim. Biophys. Acta 130, 543-546, 1986
A:Title: Covalent structure of bovine chymotrypsinogen A.
A:Reference number: A90572; MUID:67183949; PMID:5972866
A:Accession: A93158
A:Molecule type: protein
A:Residues: 1-101, N', 103-245 <MBL>
A:Note: disulfide bonds were determined
R:Cutruzzola, P.; Ascenzi, P.; Barra, D.; Bolognesi, M.; Menegatti, E.; Sarti, P.; Schneid
Biochim. Biophys. Acta 1161, 201-208, 1993
A:Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catalytic
A:Reference number: S29650; MUID:93160238; PMID:8431470
A:Accession: S29650
A:Molecule type: protein
A:Residues: 1-12,16-27,149-160,181-200 <CUT>
R:Smillie, L.B.; Hartley, B.S.
Biochem. J. 101, 232-241, 1966
A:Title: Histidine sequences in the active centres of some 'serine' proteinases.
A:Reference number: A90236; MUID:67181723; PMID:5971785
A:Accession: A90236
A:Contents: annotation; active site
R:Blktott, J.J.; Blow, D.M.; Henderson, R.; Steitz, T.A.
Philos. Trans. R. Soc. Lond. B257, 67-76, 1970
A:Title: The structure of alpha-chymotrypsin.
A:Reference number: A93754
A:Contents: annotation; X-ray crystallography
C:Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-245/Product: chymotrypsinogen #status experimental <ZYM>

```

F1:13-16-146,149-245/Product: alpha-chymotrypsin #status experimental <WPT>
 F16-238/Domain: trypsin homology <TRY>
 F1:122,42-58,136-182,191-220/Disulfide bonds: #status experimental
 F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 67.8%; Score 40; DB 1; Length 245;
 Best Local Similarity 58.3%; Pred. No. 2.5;
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 CGVPAIPXXXTG 12

DB 1 CGVPAIPQVLGG 12

RESULT 7

KYRTB chymotrypsin (EC 3.4.21.1) B precursor - rat

N:Alternate names: chymotrypsinogen B

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999

C:Accession: A22658

R:Bel1, G1.; Ostaco, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.

A:Title: Identification and sequence of a rat chymotrypsin B gene.

A:Reference number: A22658; MUID:85054881; PMID:6209274

A:Accession: A22658

A:Molecule type: DNA

A:Residues: 1-263 <BEL>

A:Cross-references: GB:M22298; NID:g203653; PIDN:AAA98732.1; PID:g203654

C:Genetics:

A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-33/Domain: propeptide #status predicted <PRO>

F:34-263/Product: chymotrypsin B #status predicted <WPT>

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 67.8%; Score 40; DB 1; Length 263;

Best Local Similarity 58.3%; Pred. No. 2.5;

Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGXPAIPXXXTG 12

DB 19 CGVPTIQVLGTG 30

RESULT 8

A21195 chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999

C:Accession: A21195

R:Canis lupus familiaris (dog); Luc, V.; Scheele, G.

A:Title: Identification of cDNA clones encoding secretory isoforms: sequence deter

A:Reference number: A21195; MUID:84170253; PMID:6584866

A:Accession: A21195

A:Molecule type: mRNA

A:Residues: 1-263 <PIN>

A:Cross-references: GB:M21173; NID:g163945; PIDN:AAA0841.1; PID:g163946

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 67.8%; Score 40; DB 2; Length 263;

Best Local Similarity 58.3%; Pred. No. 2.5;

Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 CGXPAIPXXXTG 12

Db 19 CGVPAIPQVLGG 30

RESULT 9

A31299 chymotrypsin (EC 3.4.21.1) precursor - human

C:Species: Homo sapiens (man)

C>Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 22-Jun-1999

C:Accession: A31299

R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsub

A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinog

A:Reference number: A31299; MUID:89134264; PMID:2917002

A:Accession: A31299

A:Molecule type: mRNA

A:Residues: 1-263 <TOM>

A:Cross-references: GB:M24400; NID:g181189; PIDN:AAA52128.1; PID:g181190

C:Genetics:

A:Gene: GDB:CTRB1; CTRB

A:Cross-references: GDB:119820; OMIM:118890

A:Map position: 16q23.1-16q23.1

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 67.8%; Score 40; DB 2; Length 263;

Best Local Similarity 58.3%; Pred. No. 2.5;

Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 CGXPAIPXXXTG 12

DB 19 CGVPAIPQVLGG 30

RESULT 10

C61414 chymotrypsin (EC 3.4.21.1) - slider turtle (fragment)

C:Species: Pseudemys scripta (slider)

C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999

C:Accession: C61414

R:Bhargava, A.K.; Barnard, E.A.

A:Title: Evolution in the pancreatic chymotrypsinogen series: N-terminal sequence determ

A:Reference number: A61414; MUID:76146602; PMID:4807189

A:Accession: C61414

A:Molecule type: protein

A:Residues: 1-16 <BHA>

C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 66.1%; Score 39; DB 2; Length 16;

Best Local Similarity 58.3%; Pred. No. 0.27;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGXPAIPXXXTG 12

DB 1 CGVPAIPQVLGG 12

RESULT 11

B61414 chymotrypsin (EC 3.4.21.1) - painted turtle (fragment)

C:Species: Chrysemys picta (painted turtle)

C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 07-May-1999

C:Accession: B61414

R:Bhargava, A.K.; Barnard, E.A.

A:Title: Evolution in the pancreatic chymotrypsinogen series: N-terminal sequence determ

A:Reference number: A61414; MUID:76146602; PMID:4807189

A:Accession: B61414

A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-17 <BHA>
 C:Keywords: hydrolase; serine proteinase

Query Match 66.1%; Score 39; DB 2; Length 17;
 Best Local Similarity 58.3%; Pred. No. 0.29;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGXPALPXXXTG 12
 ||||| :
 Db 1 CGPVALPVLVG 12

RESULT 12
 A82965
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: A82965
 A:Molecule type: DNA
 A:Residues: 1-542 <STO>
 A:Cross-references: GB:AE004950; GB:AE004950; NID:g9951776; PIDN:AA08840.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA5455

Query Match 62.7%; Score 37; DB 2; Length 542;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGXPALP 8
 ||||| :
 Db 420 CGTVALP 427

RESULT 13
 T07043
 A:Title: Probable epoxide hydrolase (EC 3.3.2.3) (clone EH3.1) - potato
 C:Species: Solanum tuberosum (potato)
 C:Date: Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
 A:Accession: T07043
 B:Scapleton, A.; Bertram, J.K.; Pinot, F.; Garbarino, J.E.; Rockhold, D.R.; Friedman, M.; Plant, J.G.; 251-258 1999
 A:Title: Cloning and expression of soluble epoxide hydrolase from potato.
 A:Reference number: 215681; MUID:19504665; PMID:7920715
 A:Accession: T07043
 A:Molecule type: mRNA
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-321 <STA>
 A:Cross-references: EMBL:U02494; NID:g407937; PIDN:AAA81889.1; PID:g407938
 A:Experimental source: cv. Lemhi Russet, wounded tuber
 C:Superfamily: tropinesterase
 C:Keywords: ether hydrolase

Query Match 61.0%; Score 36; DB 2; Length 321;
 Best Local Similarity 46.2%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGXPALPXXXTG 13
 ||||| :
 Db 50 CGTVALPDLRG 62

RESULT 14

A83170
 A:Title: Two component sensor kinase *exsG* [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 A:Accession: AD3170
 R:Wood, D.W.; Satubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erage, G.; Gillet, M.; Grant, C.; Guntimner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, J.; Karp, P.; Romero, R.; Zhang, S.; Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kem, I.; Ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AD3170
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-350 <KUR>
 A:Cross-references: GB:AB008687; PIDN:AA45778.1; PID:g17743513; GSPDB:GN00188
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: *exsG*
 A:Genome: plasmid

Query Match 61.0%; Score 36; DB 2; Length 350;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CGXPALPXXXTG 13
 ||||| :
 Db 295 GPFVTPSRQY 306

RESULT 15
 AH3382
 A:Title: Transcription repressor [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 A:Accession: AH3382
 R:DelVecchio, M.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Mani, M.; Goldman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc Natl Acad Sci U S A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A:Reference number: AD3252; PMID:11756888
 A:Accession: AH3382
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-373 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AA452227.1; PID:g17983011; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11046
 A:Map position: 1
 C:Superfamily: hypothetical protein H10753

Query Match 61.0%; Score 36; DB 2; Length 373;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGXPALPXXXTG 12
 ||||| :
 Db 349 CGLPLTVPSTG 360

Search completed: February 12, 2003, 10:28:49
 Job time : 5.2597 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:04:45 ; Search time 1.74627 Seconds
(without alignments)
308.768 Million cell updates/sec

Title: US-10-036-371-3

Perfect score: 59

Sequence: 1 CGXPAXPPXXTCY 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	88.1	245	1	CTRA_GADMO
2	52	88.1	263	1	CTRA_GADMO
3	40	67.8	245	1	CTRA_BOVIN
4	40	67.8	245	1	CTRA_BOVIN
5	40	67.8	263	1	CTR2_CANFA
6	40	67.8	263	1	CTRA_HUMAN
7	40	67.8	263	1	CTRA_RAT
8	38	64.4	591	1	HE_HENPU
9	38	64.4	626	1	DAILY_DROME
10	35	59.3	264	1	CTRL_HUMAN
11	35	59.3	264	1	CTRL_HUMAN
12	32	59.3	958	1	CHRD_HUMAN
13	34	57.6	265	1	SERI_DROME
14	34	57.6	282	1	PPNK_HAANI
15	34	57.6	312	1	USM1_HSV1
16	34	57.6	378	1	PEM1_PHACH
17	34	57.6	558	1	GPCI_RAT
18	33	55.9	231	1	CUT2_CABEL
19	33	55.9	246	1	CB11_LYCES
20	33	55.9	570	1	UREL_RHIME
21	33	55.9	730	1	DPOL_HPBVA
22	33	55.9	750	1	DPOL_HPBVZ
23	33	55.9	763	1	DPOL_HPBVP
24	33	55.9	832	1	DPOL_HPBVA
25	33	55.9	832	1	DPOL_HPBVP
26	33	55.9	832	1	DPOL_HPBVA
27	33	55.9	832	1	DPOL_HPBVP
28	33	55.9	843	1	DPOL_HPBVA
29	33	55.9	843	1	DPOL_HPBVP
30	33	55.9	843	1	DPOL_HPBVA
31	33	55.9	843	1	DPOL_HPBVP
32	33	55.9	843	1	DPOL_HPBVA
33	33	55.9	843	1	DPOL_HPBVP

P03159 hepatitis b
P17100 hepatitis b
P46889 escherichia
P98133 bos taurus
Q9tv36 sus scrofa
P01359 sus scrofa
Q03403 homo sapien
Q03404 mus musculus
Q09030 rattus norv
Q20257 caenorhabd
P57769 rattus norv
Q12587 candida mel

ALIGNMENTS

```

RESULT 1
CTRB_GADMO
ID_CTRB_GADMO STANDARD; PRT; 245 AA.
AC P80546;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin B (EC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Ruteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE. Jorvic caeca.
RX MEDLINE=96439045; PubMed=8841380;
RA Leth-Jarssen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M.,
RA Heerup P.;
RT "Structure of chymotrypsin variant B from Atlantic cod, Gadus
RT morhua.";
RL Biochim. Biophys. Acta 1297:49-56(1996).
RN [2]
RP SEQUENCE OF 1-12 AND 16-31.
RX TISSUE=Pyloric caeca;
RX MEDLINE=92111252; PubMed=1764912;
RA Asgeirsson B., Bjarnason J.B.;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
RT (Gadus morhua). Comparison with bovine chymotrypsin.";
RL Comp. Biochem. Physiol. 99B:327-335(1989).
CC -1- KINATIC ACTIVITY: Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC HSP; P00766; 1CHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 245
FT CDS 1 245
FT ACT_SITE 101 101
FT ACT_SITE 195 195
FT ACT_SITE 121
FT DISULFID 42 58
FT DISULFID 135 201
FT DISULFID 167 182
FT DISULFID 191 220
FT CONFLICT 9 11
OVT -> VIS (IN REF. 2).

```

```

FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
FT CONFLICT 21 21 R -> S (IN REF. 21.)
FT CONFLICT 25 25 S -> Q (IN REF. 21.)
FT CONFLICT 29 29 T -> Q (IN REF. 21.)
FT CONFLICT 44 44 S -> T (IN REF. 21.)
FT CONFLICT 46 46 S -> Y (IN REF. 21.)
SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

Query Match 88.1%; Score 52; DB 1; Length 263;
Best Local Similarity 89.2%; Pred. No. 0.0019;
Matches 9; Conservative 0; Mismatches 4; Indels 0;

QY 1 CGXRAIYDXTGY 13
DB 19 CRQPAISFVITGI 31

RESULT 3
ID CTRA_BOVIN STANDARD; PRT; 245 AA.
AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A [EC 3.4.21.1].
OS Bos taurus (Bovine),
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RP MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis. The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of chymotrypsin.";
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67181394; PubMed=5972866;
RA Meloun B., Klue I., Koscka V., Moravsek L., Prusik Z., Vanacek J.,
RA Keil B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie E.B., Hartley B.S.;
RT "Histidine sequences in the active centres of some 'serine' proteinases."
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "1. Serine proteinases. The structure of alpha-chymotrypsin.";
RL Philos. Trans. R. Soc. Lond., B. Biol. Sci. 257:67-76(1970).
RN [7]

```

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
 RA MEDLINE=70177557; PubMed=5442169;
 RX Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xiong M.H.;
 RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
 RL chymotrypsin, and implications for zymogen activation.";
 RN Biochemistry 9:1997-2009 (1970).
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
 RA MEDLINE=92078042; PubMed=6914399;
 RX Rabin G.H., Silverton E.W., Davies D.R.;
 RT "Refinement of the structure of gamma-chymotrypsin at 1.9-A resolution.
 RL Comparison with other serine proteases.";
 RN J. Mol. Biol. 148:449-479 (1981).
 RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
 RA MEDLINE=86011575; PubMed=4046030;
 RX Tsukada H., Blow D.M.;
 RT "Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
 RN J. Mol. Biol. 184:703-711 (1985).
 RL J. Mol. Biol. 184:703-711 (1985).
 CC -:- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
 Phe-|-Xaa, Leu-|-Xaa.
 CC -:- SUBCELLULAR LOCATION: Extracellular.
 CC -:- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -:- DATABASE: NMR=worthington enzyme manual;
 CC PDB: <http://www.worthington-biochem.com/manual/C/CHY.html>.
 DR PDB: 2C9S; 15-MAY-90.
 DR PDB: 2CGJ; 31-MAY-84.
 DR PDB: 2CHA; 31-MAY-84.
 DR PDB: 4CHA; 29-OCT-85.
 DR PDB: 5CHA; 16-OCT-87.
 DR PDB: 6CHA; 16-OCT-87.
 DR PDB: 1CHG; 27-JAN-84.
 DR PDB: 1CHO; 16-JUL-88.
 DR PDB: 28CH; 31-MAY-84.
 DR PDB: 36CH; 15-OCT-92.
 DR PDB: 45CH; 15-OCT-90.
 DR PDB: 55CH; 15-OCT-90.
 DR PDB: 65CH; 15-OCT-90.
 DR PDB: 76CH; 15-OCT-90.
 DR PDB: 86CH; 15-JUL-93.
 DR PDB: 1GCT; 15-OCT-91.
 DR PDB: 2GCT; 15-OCT-91.
 DR PDB: 3GCT; 15-OCT-91.
 DR PDB: 1GCG; 31-OCT-93.
 DR PDB: 1GWC; 31-OCT-93.
 DR PDB: 1GWD; 31-OCT-93.
 DR PDB: 1GCI; 30-APR-94.
 DR PDB: 1GCI; 30-APR-94.
 DR PDB: 1GCD; 22-JUN-94.
 DR PDB: 1GHA; 22-JUN-94.
 DR PDB: 1GHB; 22-JUN-94.
 DR PDB: 1GWH; 30-SEP-94.
 DR PDB: 2GMI; 01-NOV-94.
 DR PDB: 2GMT; 17-AUG-96.
 DR PDB: 1MTN; 17-AUG-96.
 DR PDB: 1AB9; 20-AUG-97.
 DR PDB: 1AFO; 17-SEP-97.
 DR PDB: 1CA0; 23-JUL-97.
 DR PDB: 1CBM; 23-JUL-97.
 DR PDB: 2VGC; 12-NOV-97.
 DR PDB: 2VGC; 12-NOV-97.
 DR PDB: 3VGC; 12-NOV-97.
 DR PDB: 4VGC; 12-NOV-97.
 DR PDB: 1HJA; 14-JAN-98.
 DR MEROPS: S01.001.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase: Serine protease; Digestion; Pancreas; Zymogen;

CHYMOTRYPSIN A, A CHAIN.
 CHYMOTRYPSIN A, B CHAIN.
 CHYMOTRYPSIN A, C CHAIN.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.

KW 10-structure. 1 13
 FT CHAIN 16 146
 FT CHAIN 149 245
 FT ACT_SITE 57 57
 FT ACT_SITE 102 102
 FT ACT_SITE 195 195
 FT DISULFID 1 122
 FT DISULFID 42 58
 FT DISULFID 136 201
 FT DISULFID 168 182
 FT DISULFID 191 220
 FT HELIX 12 15
 FT HELIX 24 25
 FT TURN 24 25
 FT TURN 28 29
 FT STRAND 30 34
 FT STRAND 40 46
 FT STRAND 51 54
 FT TURN 57 58
 FT TURN 62 63
 FT STRAND 65 68
 FT TURN 69 69
 FT STRAND 72 72
 FT TURN 73 74
 FT STRAND 81 90
 FT TURN 92 93
 FT STRAND 95 95
 FT TURN 96 99
 FT STRAND 100 100
 FT TURN 104 108
 FT STRAND 122 122
 FT TURN 126 127
 FT TURN 132 133
 FT STRAND 135 140
 FT TURN 146 148
 FT STRAND 154 154
 FT STRAND 156 163
 FT HELIX 165 168
 FT TURN 169 172
 FT HELIX 173 175
 FT STRAND 180 184
 FT TURN 192 193
 FT TURN 195 196
 FT STRAND 198 203
 FT TURN 204 205
 FT STRAND 216 218
 FT TURN 222 223
 FT STRAND 225 230
 FT HELIX 231 233
 FT TURN 234 234
 FT HELIX 235 243
 SQ SEQUENCE 245 AA; 25666 MW; 91A9F28E2F3E3142 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 245;
 Best Local Similarity 58.3%; Pred. No. 0.53;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGVPAIXPXXYG 12
 |||||
 DB 1 CGVPAIQVLSG 12

RESULT 4
 CTRB_BQVIN STANDARD; PRT; 245 AA.
 AC P00767;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsinogen B (EC 3.4.21.1).

```

OS Bos taurus (Bovine) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
(1)
RN SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RA MEDLINE=68238908; PubMed=5649671;
RX Smallie L.B., Furka A., Nagabushan N., Stevenson K.J., Parkes C.O.;
RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and
RT trypsinogen.";
RL Nature 218:343-346(1968).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-[Xaa, Trp]-[Xaa,
CC Phe]-[Xaa, Leu]-[Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- DATABASE: NAME=Northington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/c/CHY.html".
CC PIR; A00953; KYBOB.
DR HSSP; P00766; IACB.
DR MEROPS; S01152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 146 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 149 245 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 1 122 CHARGE RELAY SYSTEM.
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 198 222
FT DISULFID 199 222
SQ SEQUENCE 245 AA; 25755 MW; 678016446FP5FEBS CRC64;
Query Match 67.8%; Score 40; DB 1; Length 245;
Best Local Similarity 58.3%; Pred No. 0.53;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPALPXPXTG 12
DB 1 CGVPALQPVLGG 12

RESULT 5
ID_CTR2_CANFA STANDARD; PRT; 263 AA.
AC P04813;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedida; Canidae; Canis.
OX NCBI_TaxID=9615;
(1)
RN SEQUENCE FROM N.A.
RA MEDLINE=84170253; PubMed=6584866;
RX Pinsky S.D., Laforge K.S., Luc V., Scheele G.;
RT "Identification of cDNA clones encoding secretory isoenzyme forms:
RT sequence determination of canine pancreatic precchymotrypsinogen 2
RT mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-[Xaa, Trp]-[Xaa,

```

```

CC Phe]-[Xaa, Leu]-[Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01173; AAA30841.1; -.
DR PIR; A21195; A21195.
DR HSSP; P00766; IACB.
DR MEROPS; S01152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN 2
FT CHAIN 19 31 CHYMOTRYPSIN 2, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN 2, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN 2, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM.
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM.
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 134 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27787 MW; 2A2F449D813B3961 CRC64;
Query Match 67.8%; Score 40; DB 1; Length 263;
Best Local Similarity 58.3%; Pred No. 0.56;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPALPXPXTG 12
DB 19 CGVPALQPVLGG 30

RESULT 6
ID_CTR2_HUMAN STANDARD; PRT; 263 AA.
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
(1)
RN SEQUENCE FROM N.A.
RA TISSUE=Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
RT precchymotrypsinogen cDNA.";
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
(2)
RN SEQUENCE FROM N.A.

```

RA	Tissue:Pancreas;
BA	Strausberg R.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC	- - CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC	Phe-I-Xaa, Leu-I-Xaa.
CC	- - SUBCELLULAR LOCATION: Extracellular.
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as it is contained in a database
CC	and this statement is not removed. Usage by others than for commercial
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC	or send an email to licensee@sib-sib.ch)
CC	-----
DR	EMBL; M24400; AAA52128.1; ..
DR	EMBL; BC005385; AAH05385.1; ..
DR	PIR; A31299; A31299.
DR	HSPP; P00766; ICHG.
DR	MEROPS; S01.152; ..
DR	GeneW; HGNC:2521; CTRB1.
DR	MIM; 118890; ..
DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO01254; Ser_protease_Try.
DR	Fam1.PF00089; tryptsin; 1.
DR	SMART; SMO0224; TRYPSIN.
DR	SMART; SMO0224; TRYPSIN.
DR	PROSITE; PSS0240; TRYPSTN_D0M; 1.
DR	PROSITE; PSS0134; TRYPSTN_HIS; 1.
DR	PROSITE; PSS0135; TRYPSTN_SER; 1.
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT	SIGNAL 1 18 CHYMOTRYPSINOGEN B.
FT	CHAIN 19 263 CHYMOTRYPSIN B, A CHAIN.
FT	CHAIN 19 31 CHYMOTRYPSIN B, B CHAIN.
FT	CHAIN 34 164 CHYMOTRYPSIN B, C CHAIN.
FT	CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT	ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 150 219 BY SIMILARITY.
FT	DISULFID 186 200 BY SIMILARITY.
FT	DISULFID 209 238 BY SIMILARITY.
FT	SEQUENCE 263 AA; 27870 MW; 4C1C05A940B8701 CRC64;
SQ	SEQUENCE 263 AA; 27870 MW; 4C1C05A940B8701 CRC64;
Query Match	67.8%; Score 40; DB 1; Length 263;
Best Local Similarity	58.3%; Pred. No. 0.56;
Matches	7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY	1 CXPAIXPXXTG 12
DB	19 CGVPHPLVLSG 30
RESULT 7	
CTRB RAT	STANDARD; PRT; 263 AA.
ID	C07338
DT	01-APR-1988 (Rel. 07, Created)
DT	01-APR-1988 (Rel. 07, Last sequence update)
DE	Chymotrypsin precursor (SC 3.4.21.1)
DE	Chymotrypsin precursor (SC 3.4.21.1)
GN	CTRB1.
OC	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxId=10116;	
RA	Sequence from N.A.
RA	Medline:65054481; PubMed:6209274;
RA	Reil G.; Quiroga C.; Quiroga M.; Valenzuela P.; Craik C.S.
RA	"Sea urchin hatching enzyme (envelysin)": cDNA cloning and deprivation
RA	Nomura K., Shimizu T., Kinoh H., Sendai Y., Inomata M., Suzuki N.;
RA	J. Biol. Chem. 259:14265-14270(1984).
RA	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RA	Tissue:Pancreas;
BA	Strausberg R.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC	- - CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC	Phe-I-Xaa, Leu-I-Xaa.
CC	- - SUBCELLULAR LOCATION: Extracellular.
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as it is contained in a database
CC	and this statement is not removed. Usage by others than for commercial
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC	or send an email to licensee@sib-sib.ch)
CC	-----
DR	EMBL; M24400; AAA52128.1; ..
DR	EMBL; BC005385; AAH05385.1; ..
DR	PIR; A31299; A31299.
DR	HSPP; P00766; ICHG.
DR	MEROPS; S01.152; ..
DR	GeneW; HGNC:2521; CTRB1.
DR	MIM; 118890; ..
DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO01254; Ser_protease_Try.
DR	Fam1.PF00089; tryptsin; 1.
DR	SMART; SMO0224; TRYPSIN.
DR	SMART; SMO0224; TRYPSIN.
DR	PROSITE; PSS0240; TRYPSTN_D0M; 1.
DR	PROSITE; PSS0134; TRYPSTN_HIS; 1.
DR	PROSITE; PSS0135; TRYPSTN_SER; 1.
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT	SIGNAL 1 18 CHYMOTRYPSINOGEN B.
FT	CHAIN 19 263 CHYMOTRYPSIN B, A CHAIN.
FT	CHAIN 19 31 CHYMOTRYPSIN B, B CHAIN.
FT	CHAIN 34 164 CHYMOTRYPSIN B, C CHAIN.
FT	CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT	ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 150 219 BY SIMILARITY.
FT	DISULFID 186 200 BY SIMILARITY.
FT	DISULFID 209 238 BY SIMILARITY.
FT	SEQUENCE 263 AA; 27870 MW; 4C1C05A940B8701 CRC64;
SQ	SEQUENCE 263 AA; 27870 MW; 4C1C05A940B8701 CRC64;
Query Match	67.8%; Score 40; DB 1; Length 263;
Best Local Similarity	58.3%; Pred. No. 0.56;
Matches	7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY	1 CXPAIXPXXTG 12
DB	19 CGVPHPLVLSG 30
RESULT 7	
CTRB RAT	STANDARD; PRT; 263 AA.
ID	C07338
DT	01-APR-1988 (Rel. 07, Created)
DT	01-APR-1988 (Rel. 07, Last sequence update)
DE	Chymotrypsin precursor (SC 3.4.21.1)
DE	Chymotrypsin precursor (SC 3.4.21.1)
GN	CTRB1.
OC	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebr

[1] SEQUENCE FROM N.A.
MEDLINE=94093544; PubMed=8268911;
Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.:
"A tight cluster of five unrelated human genes on chromosome
17q22.1".
Hum. Mol. Genet. 2:1589-1595(1993)
CL - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL: X71874; CAAS0710.1; --
HSP: P05063.256;
GenPept: HGNC:2524; CTRL.
MIM: 118888;
InterPro: IPR001314; Chymotrypsin.
Pfam: PF001254; Ser protease_Try.
PRINTS: PR00722; CHYMOTRYPSIN.
SMART: SM0020; TRYD_SPC; 1.
PROSITE: PS00240; TRYPsin_DOW; 1.
PROSITE: PS00134; TRYPsin_HIS; 1.
PROSITE: PS00135; TRYPsin_SER; 1.
Hydrolase; Serine protease; Glycylserine; Zymogen; Signal.
FT SIGNAL 1 35
FT PROPEP 1 35
FT ACT_SITE 24 264
FT ACT_SITE 75 75
FT ACT_SITE 121 121
FT ACT_SITE 214 214
FT CARBOHYD 114 114
FT DISULFID 19 141
FT DISULFID 60 76
FT DISULFID 155 220
FT DISULFID 187 201
FT DISULFID 210 239
FT DISULFID 264 AA; 28002 MW; 3f62f02FA6DDB4 CRC64;
Query Match Score 59.38; Score 35; DB 1; Length 264;
Best Local Similarity 75.08; Pred. No. 6;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGXPAIXP 8
DB 19 CGTPIXP 26
RESULT 11
ID CHRD_MOUSE STANDARD; PRT; 948 AA.
AC QZ00E2; 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chordin precursor.
GN CHRD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6SJ/F1;
RA Lu B., Bachiller D., Agius E., Piccolo S., De Robertis E.M.:
"BMP-binding domains in the chordin secreted protein".
EMBL: X71874; CAAS0710.1; --

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Lin X., Mattei B., McIntosh T.C., McLeod M., McPherson D.,
McKervey G., Mishina N.Y., Morley K., Morris J., Washburn A.,
Watt S.G., Watson R.K., Wasylyk B., Murphy L., Muzny D.M., Nelson D.L.,
Palmiter R.D., Platts R.S., Pan S.S., Pollard J.R., Paciel J.E.,
Rao P.H., Pittman G.S., Pan S.S., Pollard J.R., Puri V., Reese M.G.,
Reiner B.C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shier E., Spadling A.C., Stapleton M., Strong R., Sun E.H.,
Swirski R., Tecior C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weissrock G.M., Weissenbach J., O.A.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao X.,
Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao X., Zhu X.,
Zheng X.H., Zhong N., Zhong N., Zhou X., Zhou X., Zhou X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin A.P., Whitehead J.P., Smith H.O.,
The genome project group, International Human Genome Sequencing Consortium,
Science 281:2158-2195(2000)
CL FUNCTION: HEPARAN SULFATE CELL SURFACE PROTEOGLYCAN. REQUIRED FOR
CELL DIVISION PATTERNING DURING POST-EMBRYONIC DEVELOPMENT OF THE
NERVOUS SYSTEM.
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(Potential).
CC -I- SIMILARITY: BELONGS TO THE GLYCOPCAN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL: U31985; AAA97401.1; --
EMBL: AE003554; AAF50358.1; --
FlyBase: FBgn0011577; dally.
InterPro: IPR001863; Glypican.
Pfam: PF01153; Glypican; 1.
DR PROSITE: PS01207; GLYPICAN; FALSE NEG.
KW Proteoglycan; Heparan sulfate; Glycoprotein; Signal; GPI-anchor.
FT SIGNAL 1 26
FT CHAIN 27 622
FT REMOVED IN NATURE FORM (POTENTIAL).
FT PROPEP 101 101
FT ACT_SITE 150 150
FT CARBOHYD 187 187
FT CARBOHYD 187 187
FT CARBOHYD 549 549
FT CARBOHYD 569 569
FT CARBOHYD 573 573
FT CARBOHYD 601 601
SQ SEQUENCE 626 AA; 69031 MW; 1182AD8AD0ED4D3 CRC64;
Query Match Score 64.41; Score 38; DB 1; Length 626;
Best Local Similarity 41.74; Pred. No. 3; 4;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 CGXPAIXPXTS 12
DB 406 CGTSUTPVSSG 417
RESULT 10
ID CTRL_HUMAN STANDARD; PRT; 264 AA.
AC P40313;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-).
GN CTRL (Human)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Carnivora; Felidae; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]

	RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
	RA	Zhang X.H., Zhong F.N., Zhong W., Zhu S., Zhu X., Smith H.O.,
	RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
	RL	"The genome sequence of <i>Drosophila melanogaster</i> ."
	RL	Science 287:2185-2195(2000).
	CC	-1- FUNCTION: MAJOR FUNCTION MAY BE TO AID IN DIGESTION.
	CC	-1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THE LARVAL GUT.
	CC	-1- DEVELOPMENTAL STAGE: EN ABUNDANCE THROUGHOUT THE LARVAL STAGE.
	CC	TOWARD THE END OF THE ADULT STAGE.
	CC	-1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SER1.
	CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
	CC	-1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
	CC	gene model prediction.
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation
	CC	at the European Bioinformatics Institute. There are no restrictions on its
	CC	use by non-profit institutions as long as its content is not commercial-
	CC	ized and this statement is not removed. Usage by and/or commercial-
	CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
	CC	or send an email to licenses@isb-sib.ch).
	CC	-----
	DR	EMBL; W41379; BAB02552.1; ..
	DR	EMBL; W41379; BAB02553.1; ..
	DR	EMBL; AE003771; AAF56971.1; ALT_SEQ.
	DR	EMBL; AE003771; AAF56972.1; ..
	DR	PIR; J50260; J50260.
	DR	PIR; A38894; A38894.
	DR	HSSP; P00761; LEPT.
	DR	MEROPS; S01.UPA; ..
	DR	Flybase; FBgn0003356; Ser99Da.
	DR	Flybase; FBgn0003357; Ser99Db.
	DR	InterPro; IPR001314; Chymotrypsin
	DR	InterPro; IPR001016; Ser-Protease_v8
	DR	Pfam; PF00069; Trypsin_Protease_Try.
	DR	PRINTS; PR00722; CHYMOTRYPSIN.
	DR	PRINTS; PR00839; VBPROTEASE.
	DR	SMART; SM00020; TRYSPIN_SPC; ..
	DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
	DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
	DR	PROSITE; PS00135; TRYPSIN_SER; 1.
	KW	Hydrolase; Serine protease; Signal; Zymogen; Multigene family.
	FT	SIGNAL 1 21 PROBABLE.
	FT	PROPEP 22 35
	FT	CHAIN 36 265
	FT	ACT_SITE 78 79
	FT	ACT_SITE 125 126
	FT	DISULFID 242 243
	FT	DISULFID 215 216
	FT	DISULFID 183 184
	FT	DISULFID 169 201
	FT	DISULFID 121 239
	FT	CONFLICT 14 14 A -> T (IN REF. 1; BAB02553).
	FT	SEQUENCE 265 AA; 28469 MW; 44948C523FD03384B CRC64;
	SC	-----
	Query Match	57.6%; Score 34; DB 1; Length 265;
	Best Local Similarity	50.0%; Pred. No. 9.7;
	Matches	6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
	QY	2 CXPAXIXPYTGVI 13
	DB	242 GNAPVFSVTGVI 253
	RESULT 14	-----
	PNK; HALN1	STANDARD; PRT; 282 AA.
	ID _PNK; HALN1	AD OHNK7;
	DT 16-OCT-2001 (Rel. 40, Created)	
	DT 16-OCT-2001 (Rel. 40, Last sequence update)	
	DT 16-OCT-2001 (Rel. 40, Last annotation update)	
	DE Probable inorganic polyphosphate(ATP-NAD kinase (EC 2.7.1.23))	

```

DE (Poly(P)/ATP NAD Kinase).
GN PNPK OR VNG1900C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504493; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lesky S.R., Baliga N.S., Thorson V., Shroga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Gao Y.A.,
RA Lettauver B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT Genome sequence of Halobacterium species NRC-1.
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC pyrophosphate as a source of phosphorus (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + NAD(+) = ADP + NADP(+).
CC -1- COFACTOR: Requires divalent metal ions for activity (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X02138; CAA32075.1; -
CC EMBL; X02138; CAA32075.1; -
CC EMBL; X00438; CAA25126.1; -
CC EMBL; L00036; AAA96678.1; -
CC PIR; A05242; Q0BE07.
CC InterPro: IPR000087; Collagen.
CC InterPro: IPR000714; BHV Unk.
CC Pfam; PF02053; Gene66; 1.
CC Zinc-finger.
CC FT ZN FING 271 293
CC ST ID 312 AA; 34055 MW; 12D01B0E7C920EA3 CRC64;
CC
CC Query Match 57.6%; Score 34; DB 1; Length 282;
CC Best Local Similarity 50.0%; Pred. No. 10;
CC Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
CC
CC QY 1 CCKPAIXPXTG 12
CC ||| |
CC Db 207 CGAALPFLVTG 218
CC
CC RESULT 15
CC US10_HSV1
CC ID P06166
CC AC P06166
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1998 (Rel. 06, Last sequence update)
CC DT 01-APR-1993 (Rel. 23, Last annotation update)
CC DE Varion protein US10.
CC GN US10
CC OS Herpes simplex virus (type 1 / strain 17).
CC OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC OC Alphaherpesvirinae; Simplexvirus.
CC OX NCBI_TaxID=10299;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC MEDLINE=85160822; PubMed=2984429;
CC McGeech D.J., Dolan A., Donald S., Rixon F.J.;
CC "Sequence determination and genetic content of the short unique
CC region in the genome of herpes simplex virus type 1."
CC J. Mol. Biol. 181:1-13(1985).
CC [2]
CC RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=84169548; PubMed=6324121;
RA Rixon F.J., McGeech D.J.;
RT "A 3' co-terminal family of mRNAs from the herpes simplex virus type
RT 1 short region: two overlapping reading frames encode unrelated
RT polypeptide one of which has highly reiterated amino acid sequence.";
RA Nucleic Acids Res. 12:2473-2487(1984).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 US10,
CC HSV-1 66, HSV-4 ORF3, AND VZV 64/69.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X14112; CAA32275.1; -
CC EMBL; X02138; CAA32075.1; -
CC EMBL; X00438; CAA25126.1; -
CC EMBL; L00036; AAA96678.1; -
CC PIR; A05242; Q0BE07.
CC InterPro: IPR000087; Collagen.
CC InterPro: IPR000714; BHV Unk.
CC Pfam; PF02053; Gene66; 1.
CC Zinc-finger.
CC FT ZN FING 271 293
CC ST ID 312 AA; 34055 MW; 12D01B0E7C920EA3 CRC64;
CC
CC Query Match 57.6%; Score 34; DB 1; Length 312;
CC Best Local Similarity 62.5%; Pred. No. 11;
CC Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 CCKPAIXP 8
CC ||| |
CC Db 302 CGDPAIXP 309
CC
CC Search completed: February 12, 2003, 10:23:14
CC Job time : 3.74627 secs

```

GenCore Version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:16:06 ; Search time 9.19701 Seconds
(without alignments)
291.248 Million cell updates/sec

Title: US-10-036-371-3

Perfect score: 59

Sequence: 1 CGXPAIXPXXTCY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp.archaea.*

2: sp.bacteria.*

3: sp.fungi.*

4: sp.human.*

5: sp.invertebrate.*

6: sp.mammal.*

7: sp.moll.*

8: sp.orcell.*

9: sp.phage.*

10: sp.plant.*

11: sp.todent.*

12: sp.virus.*

13: sp.vertibrate.*

14: sp.unclassified.*

15: sp.rvirus.*

16: sp.bacteriap.*

17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	88.1	23	13 Q9PR50	Q9PR50 gadus morhu
2	52	88.1	24	13 Q9PR51	Q9PR51 gadus morhu
3	52	88.1	263	13 Q9PMQ6	Q9PMQ6 gadus morhu
4	51	86.4	260	13 Q9W7Q3	Q9W7Q3 paralichthy
5	44	74.6	261	13 Q9W7Q4	Q9W7Q4 paralichthy
6	44	74.6	263	11 Q9DC86	Q9DC86 mus musculu
7	44	74.6	263	11 Q9DC88	Q9DC88 mus musculu
8	44	74.6	263	11 Q9CR35	Q9CR35 mus musculu
9	43	72.9	1365	10 Q9ZP14	Q9ZP14 volvox cart
10	38	64.4	623	5 Q966V5	Q966V5 drosophila
11	37	62.7	103	11 Q9JGF2	Q9JGF2 mus musculu
12	37	62.7	542	16 Q9HTB5	Q9HTB5 pseudomonas
13	36	61.0	264	11 Q9D960	Q9D960 mus musculu
14	36	61.0	264	11 Q9D7P8	Q9D7P8 mus musculu
15	36	61.0	264	11 Q9EQ28	Q9EQ28 rattus norv
16	36	61.0	264	11 Q9ER05	Q9ER05 mus musculu

17	36	61.0	321	10 Q41412	Q41412 solanum tub
18	36	61.0	350	16 Q8UK14	Q8UK14 acrobacteri
19	36	61.0	373	16 Q8YGV9	Q8YGV9 bruceella me
20	36	61.0	376	8 Q9XR04	Q9XR04 chlamydomon
21	36	61.0	376	8 Q9XR03	Q9XR03 chlamydomon
22	36	61.0	376	8 Q9XR02	Q9XR02 chlamydomon
23	36	61.0	376	10 Q65774	Q65774 chlamydomon
24	36	61.0	376	10 Q65777	Q65777 chloromonas
25	36	61.0	485	12 Q8QVY3	Q8QVY3 infectio
26	36	61.0	840	12 Q9E9H8	Q9E9H8 hepatitis b
27	36	61.0	840	12 Q9E9H4	Q9E9H4 hepatitis b
28	36	61.0	842	12 Q9E9H2	Q9E9H2 hepatitis b
29	36	61.0	843	12 Q9DHS2	Q9DHS2 hepatitis b
30	36	61.0	843	12 Q9DHS7	Q9DHS7 hepatitis b
31	36	61.0	843	12 Q9E9A8	Q9E9A8 hepatitis b
32	36	61.0	843	12 Q9E9A5	Q9E9A5 hepatitis b
33	36	61.0	843	12 Q9E9A1	Q9E9A1 hepatitis b
34	36	61.0	843	12 Q9E9A0	Q9E9A0 hepatitis b
35	36	61.0	843	12 Q9E9H9	Q9E9H9 hepatitis b
36	36	61.0	843	12 Q9E9H5	Q9E9H5 hepatitis b
37	36	61.0	940	13 Q57465	Q57465 gallus gall
38	36	61.0	940	13 Q9FSP2	Q9FSP2 gallus gall
39	35	59.3	40	13 Q9FSP2	Q9FSP2 gallus gall
40	35	59.3	111	17 Q57786	Q57786 pyrococcus
41	35	59.3	120	17 Q59100	Q59100 pyrococcus
42	35	59.3	189	1 Q52046	Q52046 halobacteri
43	35	59.3	333	16 Q93RX9	Q93RX9 streptomyce
44	35	59.3	351	3 Q9C0Y3	Q9C0Y3 schizosacch
45	35	59.3	385	2 Q9XZJ3	Q9XZJ3 raistonia m

ALIGNMENTS

RESULT 1

ID	Q9PR50	PRELIMINARY;	PRT;	23 AA.
AC	Q9PR50;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Chymotrypsin isoenzyme CHT2 (Fragment).			
OS	Gadus morhua (Atlantic cod).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	NCBI_TaxID=8049;			
RN	[1]			
RP	SEQUENCE			
RX	MEDLINE:96043258; PubMed:7584866;			
RA	Rae A.J., Flengrud R., Sletten K.;			
RT	"Chymotrypsin isoenzymes in Atlantic cod; differences in kinetics and			
RI	substrate specificity.," Physiol.			
RI	112B:393-398(1995).			
FT	NON_TER			
FT	NON_TER			
FT	NON_TER			
SQ	SEQUENCE	23 AA; 2356 MW; ASEC2A7B7EB3F5E CRC64;		

Query Match 88.1%; Score 52; DB 13; Length 23;
Best local similarity 69.2%; Pred. No. 0.001;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGXPAIXPXXTCY 13

Db 1 CGRPAISPVTGY 13

RESULT 2

ID	Q9PR51	PRELIMINARY;	PRT;	24 AA.
AC	Q9PR51;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Chymotrypsin isoenzyme CHT1 (P-fragment).
OS Gadus morhua (Atlantic cod).
OC Actinopterygii; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
XX [1]
RN REPEATED
RP SEQUENCE
RA MEDLINE=96043250; PubMed=7504866;
RA Raaij A.J., Flengsrud R., Sletten K.;
RT "Chymotrypsin isoenzymes in Atlantic cod; differences in kinetics and
RT substrate specificity.";
RL Comp. Biochem. Physiol. 112B:393-398(1995).
DR HSP; P00766; IACB.
FT NON_TER 1 1
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2400 MW; 0A16ACAB7B67F68D CRC64;

Query Match 88.1%; Score 52; DB 13; Length 24;
Best Local Similarity 69.2%; Pred. No. 0.0011; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 4;

Qy 1 CGXPALPXPVTGY 13
Db 1 CNFPAIQPVITGY 13

RESULT 3
ID Q9W7Q6 PRELIMINARY; PRT; 263 AA.
AC Q9W7Q6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chymotrypsin B precursor (EC 3.4.21.11).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
XX [1]
RN REPEATED
RP SEQUENCE FROM N.A.
RA TISSUE=PYLORIC CAECA;
RX MEDLINE=20464334; PubMed=11011764;
RA Spilliaert R., Gudmundsdottir A.;
RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B.";
RL Microb. Comp. Genomics 5:41-50(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AJ242521; CAB43766.1; -.
DR HSSP; P00766; ICHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYF_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Signal.
RN SIGNAL 1 16
RN CHAIN 17 263
SQ SEQUENCE 263 AA; 28175 MW; EF61B18A34BE57C CRC64;

Query Match 88.1%; Score 52; DB 13; Length 263;
Best Local Similarity 69.2%; Pred. No. 0.012; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 4;

Qy 1 CGXPALPXPVTGY 13
Db 1 CNFPAIQPVITGY 13

```

```

Db 17 CGSPALPQVITGY 29

RESULT 4
Q9W7Q3
ID Q9W7Q3 PRELIMINARY; PRT; 260 AA.
AC Q9W7Q3
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chymotrypsinogen 2.
OS Paralicthys olivaceus (Plounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
XX [1]
RN REPEATED
RP SEQUENCE FROM N.A.
RA TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029754; BAA82366.1; -.
DR HSSP; P00766; ICHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYF_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Signal.
RN SIGNAL 1 16
RN CHAIN 17 260
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F780C CRC64;

Query Match 86.4%; Score 51; DB 13; Length 260;
Best Local Similarity 69.2%; Pred. No. 0.018; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 CGXPALPXPVTGY 13
Db 16 CGSPALPVPITGY 28

RESULT 5
Q9W7Q4
ID Q9W7Q4 PRELIMINARY; PRT; 261 AA.
AC Q9W7Q4
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Plounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
XX [1]
RN REPEATED
RP SEQUENCE FROM N.A.
RA TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.

```

```

DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_DOM; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D709A9D65395B7D CRC64;

Query Match 74.6%; Score 44; DB 13; Length 261;
Best Local Similarity 66.7%; Pred No. 0.5;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPAIXPXTGY 13
DB 17 CGVPAIQPVLTG 29

RESULT 6
Q9DCB6 PRELIMINARY; PRT; 263 AA.
AC Q9DCB6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 220008D09R1K protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakata I.,
RA Saito T., Okazaki Y., Gotobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
RA Schirra L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Raski H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.*;
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC ENBL; AK007566; BAB25112.1; -.
DR HSSP; P00766; IGCT.
DR MEROPS; S01.152; -.
DR MGD; MG1:1913723; 220008D09R1K.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638F89F05A92F CRC64;

Query Match 74.6%; Score 44; DB 11; Length 263;
Best Local Similarity 66.7%; Pred No. 0.5;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPAIXPXTGT 12
DB 19 CGVPAIQPVLTG 30

```

```

Query Match 74.6%; Score 44; DB 11; Length 263;
Best Local Similarity 66.7%; Pred No. 0.5;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPAIXPXTGT 12
DB 19 CGVPAIQPVLTG 30

RESULT 7
Q9DXH8 PRELIMINARY; PRT; 263 AA.
AC Q9DXH8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 220008D09R1K protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakata I.,
RA Saito T., Okazaki Y., Gotobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
RA Schirra L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Raski H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.*;
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC ENBL; AK007566; BAB25112.1; -.
DR HSSP; P00766; IGCT.
DR MEROPS; S01.152; -.
DR MGD; MG1:1913723; 220008D09R1K.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638F89F05A92F CRC64;

Query Match 74.6%; Score 44; DB 11; Length 263;
Best Local Similarity 66.7%; Pred No. 0.5;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGXPAIXPXTGT 12
DB 19 CGVPAIQPVLTG 30

```

RESULT 8
 Q9CR35 ID Q9CR35 PRELIMINARY; PRT; 263 AA.
 AC Q9CR35;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 GN 2200008D09Rik protein.
 DN 2200008D09Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=CS7BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
 RX MEDLINE=21085660; PubMed=1117851;
 RA Kawai J., Harada A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 RA Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schrim L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Togo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Yamaoka Y., Yashida K., Yashida Y., Kawaji H., Kontaui S.,
 RA Hasekawa Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690 (2001)
 RC NATURE 409:685-690 (2001)
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
 CC TRYPsin FAMILY
 DR EMBL; AK008927; BAB25971.1; -
 DR EMBL; AK003079; BAB22553.1; -
 DR EMBL; AK007765; BAB25241.1; -
 DR EMBL; AK007815; BAB25280.1; -
 DR EMBL; AK008729; BAB25861.1; -
 DR EMBL; AK008888; BAB25954.1; -
 DR HSP; P00766; 1GCT.
 DR MEROPS; S01.152; -
 DR MGD; MG1.191373; 2200008D09Rik.
 DR InterPro; IPR001314; Chymotrypsin.
 DR Pfam; PF00089; trypsin; 1
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TVP_SPC; 1.
 DR PROSITE; PS00240; TRYPsin DOM; 1.
 DR PROSITE; PS00134; TRYPsin HIS; UNKNOWN 1.
 DR PROSITE; PS00135; TRYPsin SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match 74.6%; Score 44; DB 11; Length 263;
 Best Local Similarity 66.7%; Pred. No. 0.5;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CGXPAIXPXTG 12
 |||||
 Db 19 CGVPAIQPLTG 30
 |||||

RESULT 9
 Q9ZP14 ID Q9ZP14 PRELIMINARY; PRT; 1365 AA.
 AC Q9ZP14;
 DT 01-NOV-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 GN Alkaline phosphatase precursor [EC 3.1.3.1].
 OS Volvox carteri.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Volvocaceae; Volvox.
 NCBI_TaxID=3067;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=HK 10;
 RX MEDLINE=9908915; PubMed=9880549;
 RA Hallmann A.;
 RT "Enzymes in the extracellular matrix of Volvox: an inducible, calcium-
 dependent phosphatase with a modular composition.";
 RL J. Biol. Chem. 274:1691-1697 (1999).
 DR EMBL; AJ012458; CAA10030.1; -
 KW Hydrolase; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 1365 ALKALINE PHOSPHATASE.
 SQ SEQUENCE 1365 AA; 146283 MW; 0290DASCA4792CC4 CRC64;
 Query Match 72.9%; Score 43; DB 10; Length 1365;
 Best Local Similarity 53.8%; Pred. No. 4.1;
 Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 CGXPAIXPXTG 13
 |||||
 Db 1063 CGTATNPAAGY 1075
 |||||
 RESULT 10
 Q96EV5 ID Q96EV5 PRELIMINARY; PRT; 623 AA.
 AC Q96EV5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 GN Dally.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Euryptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RA Teuda M., Tzumi S., Nakato H.;
 RT "Transcriptional and Posttranscriptional Regulation of the dally Gene,
 a Drosophila Integral Membrane Proteoglycan";
 RL Submitted (DSC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB052367; BAB60703.1; -
 DR EMBL; AB052365; BAB60703.1; JOINED.
 DR EMBL; AB052366; BAB60703.1; JOINED.
 DR FlyBase; FBgn0011577; dally.
 DR InterPro; IPR001863; Glypican.
 DR Pfam; PF01153; Glypican; 1.
 SQ SEQUENCE 623 AA; 68791 MW; 89A8410D6BCE3411 CRC64;
 Query Match 64.4%; Score 38; DB 5; Length 623;
 Best Local Similarity 41.7%; Pred. No. 20;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 CGXPAIXPXTG 12
 |||||
 Db 403 CGTSLTPYSG 414
 |||||
 RESULT 11
 Q9JUF2 ID Q9JUF2 PRELIMINARY; PRT; 103 AA.
 AC Q9JUF2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Brain cDNA, Clone MKCD-0953, similar to Mus musculus SPAG4 protein
DE alternative spliced.
OS Mus musculus (Mouse)
OC Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Otsuka N., Kuroda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RT "Isolation of full-length cDNA clones from mouse brain cDNA library
RT made by oligo-capping method.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB041554; BAA59339.1; -.
SQ SEQUENCE 103 AA; 10590 MW; 89707EAS9ED8372 CRC64;

Query Match 62.7%; Score 37; DB 11; Length 103;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGXPALXPXTG 12
DB 59 CGEPALSPFNG 70

RESULT 12
Q9HTB5 PRELIMINARY; PRT; 542 AA.
ID Q9HTB5
AC Q9HTB5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE "Isolation of full-length cDNA clones from mouse brain cDNA library
DE made by oligo-capping method.";
OS Hypothetical protein PA5455.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]_TaxID=287;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Marriner P.,
RA Hickey M.J., Brinkman F.S.H., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Brody L.L., Coulter S.N., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Garber R.L., Gentry K.R., Kae A., Larbig K., Lim R.M.,
RA Reizer J.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Smith J.J., Saver M.H., Hancock R.E.W., Lory S., Olson M.T.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
OS Pseudomonas aeruginosa.
RX PubMed=95467300;
RL EMBL; AE004959; AAC08840.1; -.
DR InterPro; IPR001296; Glycos transf.1.
DR Pfam; PF00534; Glycos transf.1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 542 AA; 61018 MW; 537F8A7978804CAB CRC64;

Query Match 62.7%; Score 37; DB 16; Length 542;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGXPALXP 8
DB 420 CGTAPALP 427

RESULT 13
Q9D960 PRELIMINARY; PRT; 264 AA.
ID Q9D960
AC Q9D960;

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 21, Last annotation update)
DE 1810004D1581k protein.
OS Mus musculus (Mouse).
CN CTRL OR 1810004D1581K.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saito K., Iizawa M., Nishi K., Kiyosawa H., Konishi Y., Yamada I.,
RA Kikuchi K., Kawada H., Ashburner M., Hara H., Kikukawa T., Saito R.,
RA Fiedler M., Gaasterland T., Giasi C., King B., Kochiwa H.,
RA Kuhl P., Lewis S., Matsuo Y., Nikaide I., Pezole G., Quackenbush J.,
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anno H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Mashima J., Mazzarelli J.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 399:55-61 (1998).
CC -!- TRYPsin FAMILY
CC -!- TRYPsin TRYPsin BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
DR EMBL; AK007333; BAB24967.1; -.
DR HSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0020; TRYp_Spc; 1.
DR PROSITE; PSS0240; TRYpSin_D0M; 1.
DR PROSITE; PSS0134; TRYpSin_HIS; UNKNOWN_1.
DR PROSITE; PSS0135; TRYpSin_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match 61.0%; Score 36; DB 11; Length 264;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGXPALXP 8
DB 19 CGVPAIAP 26

RESULT 14
Q9D7P8 PRELIMINARY; PRT; 264 AA.
ID Q9D7P8
AC Q9D7P8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 1810004D1581k protein.
OS Mus musculus (Mouse).
CN CTRL OR 1810004D1581K.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085560; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata Y., Konno H., Adachi J., Iehii Y.,
RA Akawa T., Hara A., Fukunishi Y., Kiyosawa H., Kondo S., Yananaka I.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gajobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nishida K., King B., Kochiwa H.,
RA Schrim L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohteki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC ENBL; AB020757; 4CHA.
CC HSSP; P00766; 4CHA.
CC MEROPS; S01.256; -.
CC MCD; MGI; 885558; Ctrl.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser. Protease_Try.
CC Pfam; PF00089; trypsin_1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match 61.0%; Score 36; DB 11; Length 264;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCXPALXP 8
DB 19 CGVPAITP 26

RESULT 15
QSEQ28 PRELIMINARY; PRT; 264 AA.
AC QSEQ28,
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID:10116;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-RAT PANCREAS;
RC Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC ENBL; AB020757; BAB20287.1; -.
CC HSSP; P00766; 4CHA.
CC MEROPS; S01.256; -.

```

```

DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. Protease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;

Query Match 61.0%; Score 36; DB 11; Length 264;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCXPALXP 8
DB 19 CGVPAITP 26

Search completed: February 12, 2003, 10:27:22
Job time : 13.197 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:10:40 ; Search time 2.79403 Seconds
(without alignments)
136.898 Million cell updates/sec

Title: US-10-036-371-3

Perfect score: 59

Sequence: 1 CGPALPXXXTG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5A COMB.pcp.*
 - 2: /cgn2_6/prodata/1/iaa/5B COMB.pcp.*
 - 3: /cgn2_6/prodata/1/iaa/5C COMB.pcp.*
 - 4: /cgn2_6/prodata/1/iaa/5D COMB.pcp.*
 - 5: /cgn2_6/prodata/1/iaa/5E COMB.pcp.*
 - 6: /cgn2_6/prodata/1/iaa/5F COMB.pcp.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	59.3	855	2	US-08-938-365-3
2	35	59.3	867	2	US-08-938-365-2
3	35	59.3	854	2	US-08-938-365-1
4	34	57.6	663	4	US-09-134-001C-4074
5	34	57.6	663	4	US-09-134-001C-4074
6	33	55.9	845	1	US-09-009-953-163
7	33	55.9	845	2	US-08-416-950-11
8	33	55.9	845	2	US-08-469-830-11
9	33	55.9	845	6	5196194.17
10	32	54.2	106	1	US-08-491-976-1
11	32	54.2	106	3	US-08-476-705A-7
12	32	54.2	106	4	US-08-631-469B-20
13	32	54.2	106	4	US-09-056-868B-15
14	32	54.2	154	2	US-08-330-394A-29
15	32	54.2	156	2	US-08-330-394A-22
16	32	54.2	159	1	US-08-491-976-3
17	32	54.2	162	1	US-08-048-164A-2
18	32	54.2	162	1	US-08-460-462-2
19	32	54.2	162	1	US-08-460-457-2
20	32	54.2	162	2	US-08-460-458-2
21	32	54.2	162	2	US-08-330-394A-2
22	32	54.2	163	3	US-09-036-636-7
23	32	54.2	163	4	US-09-026-632-7
24	32	54.2	163	4	US-09-325-274-7
25	32	54.2	163	4	US-09-277-716-22
26	32	54.2	382	4	US-09-609-161B-22
27	32	54.2	382	4	US-09-609-161B-22

Sequense 387, App
Patent No. 5202417
Sequense 5, Appl
Sequense 15, Appl
Sequense 19, Appl
Sequense 10, Appl
Patent No. 5202417
Sequense 26, Appl
Sequense 26, Appl
Sequense 26, Appl
Sequense 26, Appl
Patent No. 5202417
Sequense 26, Appl
Sequense 26, Appl
Sequense 26, Appl
Sequense 26, Appl
Patent No. 5202417
Sequense 317, App
Sequense 3487, Ap
Sequense 4847, Ap
Sequense 11, Appl

ALIGNMENTS

RESULT 1
US-08-938-365-3
; Sequense 3, Application US/0898365
; Patent No. 5899909
; GENERAL INFORMATION:
; APPLICANT: Yang, Pan
; TITLE OF INVENTION: HUCHORDIN AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flak Richardson P.C.
; ADDRESS: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/938,365
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/040001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-365-3

Query Watch 59.3%; Score 35; DB 2; Length 855;
Best Local Similarity 50.0%; Prod No. 96;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGPALPXXXTG 12
DB 406 CGADLIPVQTG 417

```

RESULT 2
US-08-938-365-2
; Sequence 2, Application US/08938365
; Patent No. 5989909
; GENERAL INFORMATION:
; APPLICANT: Yang, Pan
; TITLE OF INVENTION: HUCHORDIN AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,365
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Melie-John, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/040001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-8900
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-938-365-2

Query Match 59.3%; Score 35; DB 2; Length 867;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGXPALXPXXTG 12
Db 411 CGADALIPVQTG 422

RESULT 3
US-08-749-169A-3
; Sequence 3, Application US/08749169A
; Patent No. 5846770
; GENERAL INFORMATION:
; APPLICANT: RACIE, Lisa
; APPLICANT: LAVALLIE, Edward
; APPLICANT: DEROBERTIS, Edward
; TITLE OF INVENTION: CHORDIN COMPOSITIONS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/130,032A
; FILING DATE: August 4, 1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI 5284-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-130-032A-3

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,169A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI 5284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-749-169A-3

Query Match 59.3%; Score 35; DB 2; Length 954;
Best Local Similarity 50.0%; Pred. No. 11e-02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGXPALXPXXTG 12
Db 411 CGADALIPVQTG 422

RESULT 4
US-09-130-032A-3
; Sequence 3, Application US/09130032A
; Patent No. 5986056
; GENERAL INFORMATION:
; APPLICANT: LaValle, Edward
; APPLICANT: Derobertis, Edward
; TITLE OF INVENTION: HUMAN CHORDIN COMPOSITIONS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/130,032A
; FILING DATE: August 4, 1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI 5284-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8260
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-130-032A-3

```

Query Match 59.3%; Score 35; DB 2; Length 954;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCKPA:XPXATG 12
 Db 411 CCAAL:IFVOTG 422

RESULT 5

US-09-134-001C-4074
 ; Sequence 4074; Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GTC-007
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO: 4074
 ; LENGTH: 662
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4074

Query Match 57.6%; Score 34; DB 4; Length 662;
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PAIXPXTGY 13
 Db 294 PAIGPTLSG 303

RESULT 6

US-09-009-953-163
 ; Sequence 163; Application US/09009953
 ; Patent No. 41113834
 ; GENERAL INFORMATION:
 ; APPLICANT: Satte, Alessandro
 ; TITLE OF INVENTION: Identification of Broadly
 ; NUMBER OF SEQUENCES: 274
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: Microsoft
 ; CURRENT APPLICATION NUMBER: US/09/009,953
 ; FILING DATE: 21-Jan-1998
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/036,713
 ; FILING DATE: 23-JAN-1997
 ; APPLICATION NUMBER: US 60/037,432
 ; FILING DATE: 07-FEB-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 018623-011520US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 415-576-0300

INFORMATION FOR SEQUENCE 163:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 163:
 US-09-009-953-163

Query Match 55.9%; Score 33; DB 4; Length 15;
 Best Local Similarity 62.5%; Pred. No. 4.7;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCKPA:XPX 8
 Db 2 CCKPALWP 9

RESULT 7

US-08-416-950-11
 ; Sequence 11; Application US/08416950
 ; Patent No. 5780036
 ; GENERAL INFORMATION:
 ; APPLICANT: CHISARI, Francis V.
 ; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: One Market Plaza, Steuart Street Tower
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94105-1492
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,950
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US
 ; FILING DATE:
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parmelee, Steven W.
 ; REGISTRATION NUMBER: 31,990
 ; REFERENCE/DOCKET NUMBER: 14740-2-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 467-9600
 ; TELEFAX: (415) 543-5043
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 845 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-416-950-11

Query Match 55.9%; Score 33; DB 1; Length 845;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGXPALXP 8
Db 651 CGVPALMP 658

RESULT 8
US-08-469-830-11
; Sequence 11; Application US/08469830
; Patent No. 5932244
; GENERAL INFORMATION:
; APPLICANT: CHITARI, Francis V.
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,830
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-469-830-11

Query Match 55.9%; Score 33; DB 2; Length 845;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGXPALXP 8
Db 651 CGVPALMP 658

RESULT 9
5196194-17
; Patent No. 5196194
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
; NUMBER OF SEQUENCES: 21

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,621
; FILING DATE: 7-DEC-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 17:
; LENGTH: 845
5196194-17

Query Match 55.9%; Score 33; DB 6; Length 845;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGXPALXP 8
Db 651 CGVPALMP 658

RESULT 10
US-08-491-976-1
; Sequence 1; Application US/08491976
; Patent No. 5783416
; GENERAL INFORMATION:
; APPLICANT: Thim, Lars
; APPLICANT: No. 5783416ris, Kjeld
; APPLICANT: No. 5783416ris, Fanny
; APPLICANT: Bjorn, Soren B.
; APPLICANT: Christensen, P.
; APPLICANT: Nielsen, Peter F.
; TITLE OF INVENTION: Human Spasmodic Polypeptide in
; TITLE OF INVENTION: Glycosylated Form
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5783416 No. 5783416disk of No. 5783416ch America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,976
; FILING DATE: 02-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3951.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-491-976-1

Query Match 54.3%; Score 32; DB 1; Length 106;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 CGXPAIXP 8
Db 68 CQVPGISP 75

RESULT 11
US-08-476-705A-7
; Sequence 7, Application US/08476705A
; Patent No. 6063755
; GENERAL INFORMATION:
; APPLICANT: Podolsky, Daniel K.
; TITLE OF INVENTION: INTESTINAL TREFOIL PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FARSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,705A
; FILING DATE: 07 JUN 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 00786/066004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-476-705A-7

Query Match 54.2%; Score 32; DB 3; Length 106;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGXPAIXP 8
Db 68 CQVPGISP 75

RESULT 12
US-08-476-705A-7
; Sequence 7, Application US/08631469B
; Patent No. 6221840
; GENERAL INFORMATION:
; APPLICANT: Daniel K. Podolsky
; TITLE OF INVENTION: INTESTINAL TREFOIL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FARSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/631,469B
; FILING DATE: 12 APR 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/631,469
; FILING DATE: 12 APR 1996
; APPLICATION NUMBER: 08/191,352
; FILING DATE: 02 FEB 1994
; APPLICATION NUMBER: 08/037,741
; FILING DATE: 25 MAR 1993
; APPLICATION NUMBER: 07/837,192
; FILING DATE: 13 FEB 1992
; APPLICATION NUMBER: 07/655,965
; FILING DATE: 14 FEB 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 00786/322001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200107
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-631-469B-20

Query Match 54.2%; Score 32; DB 4; Length 106;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGXPAIXP 8
Db 68 CQVPGISP 75

RESULT 13
US-09-056-868B-15
; Sequence 15, Application US/09056868B
; Patent No. 6316218
; GENERAL INFORMATION:
; APPLICANT: Podolsky, Daniel K.
; TITLE OF INVENTION: INTESTINAL TREFOIL PROTEINS
; FILE REFERENCE: 00786-066005
; CURRENT APPLICATION NUMBER: US/09/056,868B
; FILING DATE: 1996-01-27
; PRIOR APPLICATION NUMBER: 08/08/476,705
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/191,352
; PRIOR FILING DATE: 1994-02-02
; PRIOR APPLICATION NUMBER: US 08/037,741
; PRIOR FILING DATE: 1993-03-25
; PRIOR APPLICATION NUMBER: US 07/837,192
; PRIOR FILING DATE: 1992-02-13
; PRIOR APPLICATION NUMBER: US 07/655,965
; PRIOR FILING DATE: 1991-02-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FARSEQ for Windows Version 4.0
; SEQ ID NO 15
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated peptide
US-09-056-868B-15

Query Match 54.2%; Score 32; DB 4; Length 106;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 CQYPAIXP 8
DB 68 CQYFOISP 75

RESULT 14
US-08-330-394A-29
; Sequence 29, Application US/08330394A
; Patent No. 5856201
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Yosef, Karmey
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: METHODS OF DETECTION USING THE
; NUMBER OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,394A
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA: CIP OF PCT/US94/04132
; APPLICATION NUMBER: 14-APR-1994
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REFERENCE/DOCKET NUMBER: 7809-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ NO. 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 154 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-330-394A-29

Query Match 54.2%; Score 32; DB 2; Length 154;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 PAIXPXXTGY 13
DB 136 PVVNPKVITY 145

RESULT 15
US-08-330-394A-22
; Sequence 22, Application US/08330394A
; Patent No. 5856201
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Yosef, Karmey
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: METHODS OF DETECTION USING THE
; NUMBER OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,394A
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA: CIP OF PCT/US94/04132
; APPLICATION NUMBER: 14-APR-1994
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REFERENCE/DOCKET NUMBER: 7809-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 156 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-330-394A-22

Query Match 54.2%; Score 32; DB 2; Length 156;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 PAIXPXXTGY 13
DB 136 PVVNPKVITY 145

Search completed: February 12, 2003, 10:30:08
Job time : 4.79403 secs


```

QY 1 CGVPAIXPXTGT 13
   |||||
Db 1 CGVPAIXPXTGT 13

RESULT 2
US-10-036-371-5
; Sequence 5, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Bovine sp.
US-10-036-371-5

Query Match 67.8%; Score 40; DB 12; Length 13;
Best Local Similarity 58.3%; Pred. No. 0.067;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGVPAIXPXTGT 12
   |||||
Db 1 CGVPAIQPVLSG 12

RESULT 3
US-09-925-297-643
; Sequence 643, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: PCT/US00/05989
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 643
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (133)

```

```

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-643

Query Match 67.8%; Score 40; DB 10; Length 146;
Best Local Similarity 58.3%; Pred. No. 0.72;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGVPAIXPXTGT 12
   |||||
Db 23 CGVPAIHPVLGS 34

RESULT 4
US-09-888-615-96
; Sequence 96, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CHENPESL, SEAN
; APPLICANT: CHANTZAKIS, GLEN
; APPLICANT: MANNING, GERALD
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-96

Query Match 67.8%; Score 40; DB 10; Length 263;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGVPAIXPXTGT 12
   |||||
Db 19 CGVPAIHPVLGS 30

RESULT 5
US-09-925-297-529
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 529
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-529

```


Wed Feb 12 11:59:35 2003

```

/ PRIOR FILING DATE: December 22, 1998
/ PRIOR APPLICATION NUMBER: 60/146,222
/ PRIOR FILING DATE: July 28, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/19330
/ PRIOR FILING DATE: September 16, 1998
/ PRIOR APPLICATION NUMBER: PCT/US98/25108
/ PRIOR FILING DATE: December 1, 1998
/ PRIOR APPLICATION NUMBER: 09/216,021
/ PRIOR FILING DATE: December 16, 1998
/ PRIOR APPLICATION NUMBER: 09/218,517
/ PRIOR FILING DATE: December 22, 1998
/ PRIOR APPLICATION NUMBER: 09/254,311
/ PRIOR FILING DATE: March 3, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/12252
/ PRIOR FILING DATE: June 22, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: September 15, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28409
/ PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: December 1, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: December 16, 1999
/ PRIOR APPLICATION NUMBER: PCT/US00/03565
/ PRIOR FILING DATE: February 11, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: February 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/05841
/ PRIOR FILING DATE: March 2, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/08439
/ PRIOR FILING DATE: March 30, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/14042
/ PRIOR FILING DATE: May 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/20710
/ PRIOR FILING DATE: July 28, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/36678
/ PRIOR FILING DATE: December 1, 2000
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: February 28, 2001
/ NUMBER OF SEQ ID NOS: 120
/ SEQ ID NO 7
/ LENGTH: 954
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-09-944-413-7
Query Match 59.3%; Score 35; DB 9; Length 954;
Best Local Similarity 50.0%; Pred No. 42;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 CCGATYKXG 12
Db 411 CCGAALFPQTG 422
RESULT 9
US-09-944-403-7
/ Sequence 7, Application US/09944403
/ Patent No. US20020165143A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin
/ APPLICANT: Botstein, David
/ APPLICANT: Eaton, Dan
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Garritsen, Mary
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul
/ APPLICANT: Grimaldi, Christopher
/ APPLICANT: Gurney, Austen
/ APPLICANT: Hillan, Kenneth

```

```

/ APPLICANT: Kljavin, Ivar
/ APPLICANT: Napier, Mary
/ APPLICANT: Roy, Margaret
/ APPLICANT: Tuma, Daniel
/ APPLICANT: Wood, William
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P2548P1C1
/ CURRENT APPLICATION NUMBER: US/09/944,403
/ CURRENT FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 09/866,028
/ PRIOR FILING DATE: 2001-05-25
/ PRIOR APPLICATION NUMBER: 60/067,411
/ PRIOR FILING DATE: December 3, 1997
/ PRIOR APPLICATION NUMBER: 60/069,334
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,335
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,278
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,425
/ PRIOR FILING DATE: December 12, 1997
/ PRIOR APPLICATION NUMBER: 60/069,666
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,694
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,602
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,897
/ PRIOR FILING DATE: December 17, 1997
/ PRIOR APPLICATION NUMBER: 60/069,873
/ PRIOR FILING DATE: December 18, 1997
/ PRIOR APPLICATION NUMBER: 60/070,440
/ PRIOR FILING DATE: January 5, 1998
/ PRIOR APPLICATION NUMBER: 60/074,086
/ PRIOR FILING DATE: February 9, 1998
/ PRIOR APPLICATION NUMBER: 60/074,092
/ PRIOR FILING DATE: February 9, 1998
/ PRIOR APPLICATION NUMBER: 60/075,945
/ PRIOR FILING DATE: February 25, 1998
/ PRIOR APPLICATION NUMBER: 60/112,850
/ PRIOR FILING DATE: December 16, 1998
/ PRIOR APPLICATION NUMBER: 60/113,296
/ PRIOR FILING DATE: December 22, 1998
/ PRIOR APPLICATION NUMBER: 60/146,222
/ PRIOR FILING DATE: July 28, 1999
/ PRIOR APPLICATION NUMBER: PCT/US98/19330
/ PRIOR FILING DATE: September 16, 1998
/ PRIOR APPLICATION NUMBER: PCT/US98/25108
/ PRIOR FILING DATE: December 1, 1998
/ PRIOR APPLICATION NUMBER: 09/216,021
/ PRIOR FILING DATE: December 16, 1998
/ PRIOR APPLICATION NUMBER: 09/218,517
/ PRIOR FILING DATE: December 22, 1998
/ PRIOR APPLICATION NUMBER: 09/254,311
/ PRIOR FILING DATE: March 3, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/12252
/ PRIOR FILING DATE: June 22, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: September 15, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28409
/ PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: December 1, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: December 16, 1999
/ PRIOR APPLICATION NUMBER: PCT/US00/03565
/ PRIOR FILING DATE: February 11, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: February 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/05841
/ PRIOR FILING DATE: March 2, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/08439
/ PRIOR FILING DATE: March 30, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/14042
/ PRIOR FILING DATE: May 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/20710
/ PRIOR FILING DATE: July 28, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/36678
/ PRIOR FILING DATE: December 1, 2000
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: February 28, 2001
/ NUMBER OF SEQ ID NOS: 120
/ SEQ ID NO 7
/ LENGTH: 954
/ TYPE: PRT
/ ORGANISM: Homo Sapien

```

```

/ PRIOR FILING DATE: February 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/05841
/ PRIOR FILING DATE: March 2, 2000
/ PRIOR FILING DATE: March 2, 2000
/ PRIOR FILING DATE: March 30, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/14042
/ PRIOR FILING DATE: May 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/20710
/ PRIOR FILING DATE: July 28, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/32678
/ PRIOR FILING DATE: December 1, 2000
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: February 28, 2001
/ NUMBER OF SEQ ID NOS: 120
/ SEQ ID NO 7
/ LENGTH: 954
/ TYPE: PRT
/ ORGANISM: Homo Sapien
/ US-09-944-403-7

```

```

Query Match          50.3%; Score 35; DB 9; Length 954;
Best Local Similarity 50.0%; Pref. No. 42;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CGXPALXPXXTG 12
    |||:|:|
Db 411 CGADALIPVQTG 422

```

RESULT 10

```

/ US-09-944-896-7
/ Sequence 7, Application US/09944896
/ Patent No. US20020168715A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin
/ APPLICANT: Botstein, David
/ APPLICANT: Perlmutter, Napoleon
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gerritsen, Mary
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul
/ APPLICANT: Grimaldi, Christopher
/ APPLICANT: Gurney, Austin
/ APPLICANT: Hillan, Kenneth
/ APPLICANT: K. Javin, Ivar
/ APPLICANT: Napier, Mary
/ APPLICANT: Roy, Margaret
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Wood, William
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE INFORMATION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: 168715A1
/ CURRENT FILING DATE: 2001-08-21
/ PRIOR FILING DATE: 2001-08-21
/ PRIOR APPLICATION NUMBER: US/09/944,896
/ PRIOR FILING DATE: 2001-05-25
/ PRIOR APPLICATION NUMBER: 60/069,334
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,335
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,278
/ PRIOR FILING DATE: December 11, 1997
/ PRIOR APPLICATION NUMBER: 60/069,425
/ PRIOR FILING DATE: December 12, 1997
/ PRIOR APPLICATION NUMBER: 60/069,696
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,694
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,702
/ PRIOR FILING DATE: December 16, 1997
/ PRIOR APPLICATION NUMBER: 60/069,870
/ PRIOR FILING DATE: December 17, 1997

```

```

/ PRIOR APPLICATION NUMBER: 60/069,873
/ PRIOR FILING DATE: December 17, 1997
/ PRIOR APPLICATION NUMBER: 60/069,017
/ PRIOR FILING DATE: December 18, 1997
/ PRIOR APPLICATION NUMBER: 60/070,440
/ PRIOR FILING DATE: January 5, 1998
/ PRIOR APPLICATION NUMBER: 60/074,086
/ PRIOR FILING DATE: February 9, 1998
/ PRIOR APPLICATION NUMBER: 60/074,092
/ PRIOR FILING DATE: February 9, 1998
/ PRIOR APPLICATION NUMBER: 60/075,945
/ PRIOR FILING DATE: February 25, 1998
/ PRIOR APPLICATION NUMBER: 60/112,650
/ PRIOR FILING DATE: December 11, 1998
/ PRIOR APPLICATION NUMBER: 60/113,288
/ PRIOR FILING DATE: December 21, 1998
/ PRIOR APPLICATION NUMBER: 60/146,222
/ PRIOR FILING DATE: July 28, 1999
/ PRIOR APPLICATION NUMBER: PCT/US98/19330
/ PRIOR FILING DATE: September 16, 1998
/ PRIOR APPLICATION NUMBER: PCT/US96/25108
/ PRIOR FILING DATE: December 1, 1998
/ PRIOR APPLICATION NUMBER: 09/216,021
/ PRIOR FILING DATE: December 16, 1998
/ PRIOR APPLICATION NUMBER: 09/218,517
/ PRIOR FILING DATE: December 22, 1998
/ PRIOR APPLICATION NUMBER: 09/254,311
/ PRIOR FILING DATE: March 3, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/12252
/ PRIOR FILING DATE: June 22, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: September 15, 1999
/ PRIOR APPLICATION NUMBER: PCT/US94/28409
/ PRIOR FILING DATE: No. US20020168715A1 September 30, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: No. US20020168715A1 September 30, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/28301
/ PRIOR FILING DATE: December 1, 1999
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: December 16, 1999
/ PRIOR APPLICATION NUMBER: PCT/US00/03565
/ PRIOR FILING DATE: February 11, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: February 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/05841
/ PRIOR FILING DATE: March 2, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/08439
/ PRIOR FILING DATE: March 30, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/14042
/ PRIOR FILING DATE: May 22, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/20710
/ PRIOR FILING DATE: July 28, 2000
/ PRIOR APPLICATION NUMBER: PCT/US00/32678
/ PRIOR FILING DATE: December 1, 2000
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: February 28, 2001
/ NUMBER OF SEQ ID NOS: 120
/ SEQ ID NO 7
/ LENGTH: 954
/ TYPE: PRT
/ ORGANISM: Homo Sapien
/ US-09-944-896-7

```

```

Query Match          50.3%; Score 35; DB 9; Length 954;
Best Local Similarity 50.0%; Pref. No. 42;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CGXPALXPXXTG 12
    |||:|:|
Db 411 CGADALIPVQTG 422

```

RESULT 11

```

US-09-944-944-7
; Sequence 7, Application US/09944944
; Patent No. US20020173463A1
; GENERAL INFORMATION:
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P2548P1C1
; CURRENT FILING DATE: 2001-09-26
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/067,411
; PRIOR FILING DATE: December 3, 1997
; PRIOR APPLICATION NUMBER: 60/069,334
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,278
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,425
; PRIOR FILING DATE: December 12, 1997
; PRIOR APPLICATION NUMBER: 60/069,696
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,694
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,702
; PRIOR FILING DATE: December 16, 1997
; PRIOR APPLICATION NUMBER: 60/069,870
; PRIOR FILING DATE: December 11, 1997
; PRIOR APPLICATION NUMBER: 60/069,873
; PRIOR FILING DATE: December 17, 1997
; PRIOR APPLICATION NUMBER: 60/068,017
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/070,440
; PRIOR FILING DATE: January 5, 1998
; PRIOR APPLICATION NUMBER: 60/074,086
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/074,092
; PRIOR FILING DATE: February 9, 1998
; PRIOR APPLICATION NUMBER: 60/075,945
; PRIOR FILING DATE: February 25, 1998
; PRIOR APPLICATION NUMBER: 60/112,850
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 60/113,296
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 60/146,222
; PRIOR FILING DATE: July 28, 1999
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: September 16, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: December 1, 1998
; PRIOR APPLICATION NUMBER: 09/216,021
; PRIOR FILING DATE: December 16, 1998
; PRIOR APPLICATION NUMBER: 09/218,517
; PRIOR FILING DATE: December 22, 1998
; PRIOR APPLICATION NUMBER: 09/254,311
; PRIOR FILING DATE: March 3, 1999

```

```

; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: June 22, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: September 15, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28409
; PRIOR FILING DATE: NO. US20020173463A1ember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: NO. US20020173463A1ember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 7
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-944-7

Query Match 59.3%; Score 35; DB 9; Length 954;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Q/ 1 CGXPALXPXTG 12
Db 411 CGADALIPVQTG 422

RESULT 12
US-09-944-907-7
; Sequence 7, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein,David
; APPLICANT: Eaton,Dan
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul
; APPLICANT: Grimaldi,Christopher
; APPLICANT: Gurney,Austin
; APPLICANT: Hillan,Kenneth
; APPLICANT: Kljavin,Ivar
; APPLICANT: Napier,Mary
; APPLICANT: Roy,Margaret
; APPLICANT: Tumas,Daniel
; APPLICANT: Wood,William
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120

```

```

; SEQ ID NO 7
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-907-7

Query Match
Best Local Similarity 59.3%; Score 35; DB 9; Length 954;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CXPALPXXXTG 12
Db 411 CGADALIPVQTG 422

RESULT 13
US-09-944-929-7
; Sequence 7, Application US/09944929
; Publication: US0020197612A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavini, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
; APPLICANT: Tomas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P2548P1C1
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US/09/944,929
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 7
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-929-7

Query Match
Best Local Similarity 59.3%; Score 35; DB 10; Length 954;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CXPALPXXXTG 12
Db 411 CGADALIPVQTG 422

RESULT 15
US-09-944-449-7
; Sequence 7, Application US/09944449
; Patent No. US200202647A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavini, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
; APPLICANT: Tomas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P2548P1C1
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/944,449
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 7
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-944-929-7

Query Match
Best Local Similarity 59.3%; Score 35; DB 9; Length 954;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CXPALPXXXTG 12
Db 411 CGADALIPVQTG 422

RESULT 14
US-09-866-028-7
; Sequence 7, Application US/09866028
; Publication: US0020059109A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher

```

Search completed: February 12, 2003, 10:31:05
Job time : 2.9791 secs

PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,877
PRIOR FILING DATE: December 13, 1997
PRIOR APPLICATION NUMBER: 60/070,940
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/071,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020102647A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020102647A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 7
LENGTH: 954
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-449-7

Query Match 59.3%; Score 35; DB 10; Length 954;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGXPAIXPXTG 12

DB 411 CGADALIEVQTG 422

..

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:01:55 ; Search time 13.4328 Seconds
(without alignments)
138.395 Million cell updates/sec

Title: US-10-036-371-4

Perfect score: 100

Sequence: 1 INVGEAVPHXMQVSLQX 20

Scoring table:

BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A: Geneseq_1010021.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.0	20	22	AA131576	Amino terminal seq
2	85.0	20	22	AA131578	Amino terminal seq
3	85.0	31	17	AAW02588	PX, porcine pancre
4	85.0	242	7	AA160061	Sequence of human
5	85.0	243	21	AA199596	Bovine chymotryps
6	85.0	243	21	AA199596	Porcine chymotryps
7	82.0	20	20	AA193621	Bovine chymotryps
8	82.0	20	20	AA193621	N-terminal of a bo
9	82.0	20	22	AA197945	N-terminal of bovi
10	82.0	146	21	AA154191	Human pancreatic c

11	82	82.0	231	22	AA198504	Human chymotrypsin
12	82	82.0	263	23	AA198278	Amino acid sequenc
13	82	80.0	242	7	AA160060	Sequence of human
14	80	80.0	242	7	AA160060	Mouse serine prote
15	80	80.0	270	6	AA161711	Human pancreas ela
16	80	80.0	270	23	AA198798	Human pancreatic t
17	79	79.0	242	13	AA1982510	Human pancreatic c
18	78	78.0	29	16	AA1984270	Human pancreatic c
19	78	78.0	268	17	AA1980682	Human pancreatic c
20	78	78.0	273	21	AA1954243	Human pancreatic c
21	77	77.0	268	17	AA1980683	Human pancreatic c
22	77	77.0	268	17	AA1988481	Human elastase IV
23	77	77.0	268	19	AA19840530	Human elastase hom
24	77	77.0	268	21	AA1989410	Human elastase IV
25	77	77.0	268	21	AA1989410	Human elastase IV
26	76	76.0	123	20	AA116550	Human 5' EST seque
27	76	74.0	195	20	AA116549	Human 5' EST seque
28	74	74.0	422	7	AA160058	Sequence of human
29	74	74.0	269	7	AA160062	Human elastase I.
30	74	74.0	269	7	AA161076	Human elastase I.
31	74	74.0	269	21	AA1954124	Human pancreatic c
32	74	74.0	279	22	AA1980513	Human pancreatic c
33	74	74.0	279	22	AA1980513	Human pancreatic c
34	71	71.0	269	8	AA170758	Human pancreatic c
35	71	71.0	269	8	AA170758	Human pancreatic c
36	70	70.0	192	21	AA1954077	Human pancreatic c
37	70	70.0	264	21	AA111710	Human serine prote
38	70	70.0	913	22	AA1982426	Novel human diago
39	70	70.0	970	23	AA1982743	Novel human diago
40	70	70.0	1031	23	AA1980993	Novel human protei
41	70	70.0	1576	22	AA19807870	Novel human diago
42	70	70.0	1576	22	AA19807870	Novel human diago
43	70	70.0	1576	22	AA19807870	Novel human diago
44	70	70.0	1576	22	AA19807870	Novel human diago
45	69	69.0	113	21	AA1954419	Human pancreatic c

ALIGNMENTS

RESULT 1

AA131576

ID AA131576 standard; peptide; 20 AA.

XX AA131576;

XX AA131576;

DT 20-APR-2001 (first entry)

XX Amino terminal sequence of cod trypsin B chain.

XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;

XX arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;

XX rheumatoid arthritis; juvenile rheumatoid arthritis; rheumatoid arthritis;

XX systemic lupus erythematosus; phlebotis; tendinitis; rashes; myalgia;

XX acne; eczema; facial seborrheic eczema; folliculitis; folliculitis; folliculitis;

XX athlete's foot; fistulae infection; ulcer; navel infection; wound;

XX scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;

XX fungal infection; autoimmune disease.

XX Gadus callarias.

XX Key

XX Location/Qualifiers

FT Key

FT Misc-difference 1

FT /label= Ser, Thr

FT Misc-difference 1

FT /label= Ser, Pro, Tyr

FT Misc-difference 20

FT /label= Asp, Gln

FT WO200078312-A2.

PN N-terminal of a bo

XX 28-DEC-2000.

XX

PF 15-JUN-2000; 2000WO-IS00005.
 PR 18-JUN-1999; 99IS-0005086.
 PA (BJAR/) BJARNASON J B.
 XX BJarnason JB;
 XX WPI; 2001-091493/10.
 XX Fish serine proteinase, useful as a cosmetic, medicament for treating
 PT eczema, psoriasis, arthritis, and in the manufacture of the medicament
 PT for treating, preventing pathogenic diseases involving receptor
 PT mediated binding -
 XX Disclosure; Page 5; 38pp; English.
 XX The specification describes a fish serine proteinase. The proteinases
 CC are useful as medicaments, for treating and preventing a disease in a
 CC human or an animal such as pain, acute inflammation, chronic arthritis,
 CC inflammation, arthritis, inflamed joints, bursitis, osteoarthritis,
 CC rheumatoid arthritis, juvenile rheumatoid arthritis, septic arthritis,
 CC fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, rash,
 CC psoriasis, acne, eczema, facial seborrheic eczema, eczema of the hands,
 CC face or neck, foreskin infections, athlete's foot, fistulae infections,
 CC infected topical ulcers, navel infections in newborns, wrinkles, scars,
 CC keloids, boils, warts and allergic itch, hemorrhoids, wounds, wound
 CC infections, wounds from burns, fungal infections and immunological and
 CC autoimmune diseases. They are also useful for removing dead or peeling
 CC skin from otherwise healthy skin, and for treating or preventing a
 CC disease in which pathogenesis is caused by bacteria, virus, fungus,
 CC parasite or a protozoan or a receptor mediated binding is involved.
 CC The present sequence represents the amino terminal of cod trypsin B
 CC chain.
 XX SQ Sequence 20 AA;
 Query Match 95.0%; Score 95; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.4e-09; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0;
 Qy 1 IVNGEEAVPHXKMQVSLQ 19
 Db 1 IVNGEEAVPHXKMQVSLQ 19
 RESULT 2
 AAB31578
 ID AAB31578 standard; peptide; 20 AA.
 AC AAB31578;
 XX 20-APR-2001 (first entry)
 DT Amino terminal sequence of bovine trypsin B chain.
 DE Fish; serine proteinase; pain; acute inflammation; chronic inflammation;
 XX arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;
 KW rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
 KW systemic lupus erythematosus; phlebitis; tendinitis; rash; psoriasis;
 KW acne; eczema; facial seborrheic eczema; foreskin infection;
 KW athlete's foot; fistulae infection; ulcer; navel infection; wrinkle;
 KW scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;
 KW fungal infection; autoimmune disease.
 XX OS Bos sp.
 XX WO200078332-A2.
 PN 28-DEC-2000.
 PD 15-JUN-2000; 2000WO-IS00005.
 XX

PR 18-JUN-1999; 99IS-0005086.
 XX (BJAR/) BJARNASON J B.
 PA BJarnason JB;
 XX WPI; 2001-091493/10.
 XX Fish serine proteinase, useful as a cosmetic, medicament for treating
 PT eczema, psoriasis, arthritis, and in the manufacture of the medicament
 PT for treating, preventing pathogenic diseases involving receptor
 PT mediated binding -
 XX Disclosure; Page 5; 38pp; English.
 XX The specification describes a fish serine proteinase. The proteinases
 CC are useful as medicaments, for treating and preventing a disease in a
 CC human or an animal such as pain, acute inflammation, chronic arthritis,
 CC inflammation, arthritis, inflamed joints, bursitis, osteoarthritis,
 CC rheumatoid arthritis, juvenile rheumatoid arthritis, septic arthritis,
 CC fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, rash,
 CC psoriasis, acne, eczema, facial seborrheic eczema, eczema of the hands,
 CC face or neck, foreskin infections, athlete's foot, fistulae infections,
 CC infected topical ulcers, navel infections in newborns, wrinkles, scars,
 CC keloids, boils, warts and allergic itch, hemorrhoids, wounds, wound
 CC infections, wounds from burns, fungal infections and immunological and
 CC autoimmune diseases. They are also useful for removing dead or peeling
 CC skin from otherwise healthy skin, and for treating or preventing a
 CC disease in which pathogenesis is caused by bacteria, virus, fungus,
 CC parasite or a protozoan or a receptor mediated binding is involved.
 CC The present sequence represents the amino terminal of bovine trypsin B
 CC chain.
 XX SQ Sequence 20 AA;
 Query Match 85.0%; Score 85; DB 22; Length 20;
 Best Local Similarity 84.2%; Pred. No. 1.6e-07; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 3;
 Qy 1 IVNGEEAVPHXKMQVSLQ 19
 Db 1 IVNGEEAVPGSNPQVSLQ 19
 RESULT 3
 AAM02588
 ID AAM02588 standard; peptide; 31 AA.
 AC AAM02588;
 XX 12-FEB-1997 (first entry)
 DT PX, porcine pancreas-derived factor N-terminal peptide.
 DE Pig; bone; calcium; resorption inhibition; formation; osteoporosis;
 KW Paget's disease; fracture repair; bone defect; osteopetrosis;
 KW metastasis; osteoblast; osteoclast.
 XX Sus scrofa.
 OS WO9619501-A1.
 PN 27-JUN-1996.
 PD 19-DEC-1995; 95WO-US16826.
 PP 19-DEC-1994; 94US-0363092.
 PR (OSTE-) OSTEOSA INC.
 PA Izbicka E, Mundy GR, Yoneda T;
 XX WPI; 1996-309523/31.
 XX

XX New isolated pancreatic-derived factor - which inhibits bone
PT resorption and increases bone formation, used to develop prods. for
PT diagnosis and therapy
XX
XS Claim 1; Fig 14; 53pp; English.
CC AA02588 is the N-terminal portion of a factor derived from pig
CC pancreas. This factor was designated PX. PX inhibits the
CC action of osteoclasts and addresses the action of osteoblasts.
CC hence PX inhibits bone resorption and promotes bone formation.
CC new factor also stimulates the proliferation of osteoblasts.
CC osteosarcoma cells (osteoblast-like cells). The factor is useful for
CC treating bone loss, e.g. due to osteoporosis or Paget's disease. The
CC factor may also be used to hasten bone fracture repair. Antibodies
CC and antagonists of PX can be used to treat osteoporosis and
CC osteoblastic metastasis
XX
SQ Sequence 31 AA;
Query Match 85.0%; Score 85; DB 17; Length 31;
Best Local Similarity 73.7%; Pred. No. 2,7e-07;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 IVNGEAVPHKXKXQVSLQ 19
Db 1 VNGEDAVPYSPKQVSLQ 19
:|||||:|||||
RESULT 4
AAP60061
ID AAP60061 standard; Protein; 242 AA.
XX
AC AAP60061;
XX
XX 06-SEP-1991 (first entry)
DT
XX
XX Enzyme; serum lipoprotein metabolism.
KW
XX Homo sapiens.
XX
XX EP198645-A.
XX
XX 22-OCT-1986.
XX
XX 07-APR-1986; 86EP-0302557.
XX
XX 23-OCT-1985; 85JP-0236686.
XX 05-APR-1985; 85JP-0072108.
XX 27-APR-1985; 85JP-0091986.
XX 26-JUL-1985; 85JP-0163964.
XX 02-DEC-1985; 85JP-0271128.
XX (SANY) SANKYO KK.
XX
XX Takiguchi Y, Tani T, Kawashima I, Erukawa H, Ohmine T;
PI Ohsuna J;
XX
XX WPI; 1986-280300/43.
DR
DR N-PSDB; AAN60075.
XX
XX New genetically engineered human pancreatic elastase - obdr.
PT using hosts modified DNA coding for enzyme
XX
XX Claim 2; Page 8; 45pp; English.
XX
XX The genetically engineered prod. can eliminate the dependency on
XX human pancreas supplies for the elastase, and avoids antibody
XX formation and possibility of anaphylaxis using porcine elastase.
XX
XX Sequence 242 AA;
SQ

Query Match 85.0%; Score 85; DB 7; Length 242;
Best Local Similarity 73.7%; Pred. No. 2,7e-06;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 IVNGEAVPHKXKXQVSLQ 19
Db 1 VNGEDAVPYSPKQVSLQ 19
:|||||:|||||
RESULT 5
AAD99596
ID AAD99596 standard; protein; 245 AA.
XX
XX AAD99596;
XX
XX 13-SEP-2000 (first entry)
DT
XX
XX Bovine chymotrypsinogen A.
XX
XX Bovine; plasminogen activator; cardiant; thrombolytic;
KW heart attack; stroke; blood clotting disorder.
XX
XX Bos taurus.
XX
XX WO200032759-A1.
PN
XX 08-JUN-2000.
PD
XX
XX 06-MAY-1999; 99WO-US09991.
XX
XX 02-DEC-1998; 98US-0110588.
XX
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
XX Lin X, Zhang XC, Tang JUN;
PI
XX WPI; 2000-422975/36.
DR
XX
XX Polypeptide with plasminogen activator activity useful as thrombolytic
PT agent for treating blood clot disorders e.g. heart attack, comprises 10
PT amino acid peptide fragment for recognition or activation of
PT plasminogen -
XX
XX Disclosure; Page 35-36; 41pp; English.
XX
XX The present sequence is bovine chymotrypsinogen. It was included in
XX a patent sequence homologues of several plasminogen activators.
XX Plasminogen is a soluble enzyme that converts plasminogen into
XX extracellular fluids of vertebrates. Plasminogen is associated with a
XX range of physiological and pathological processes. Plasminogen expression is
XX implicated in pericellular proteolysis associated with the range of
XX physiological and pathological processes. Plasminogen expression is
XX regulated by plasminogen activators which hydrolyse a peptide bond in
XX plasminogen to convert it to plasmin or form tight binding complexes
XX with plasminogen to spontaneously convert it to plasmin. The sequence
XX homology analysis has identified a six amino acid peptide involved in
XX plasminogen activation. This peptide is particularly useful when inserted
XX between amino acid residues 644 and 645 of full length human
XX plasminogen. Novel plasminogen activators have been made based upon the
XX plasminogen activation/recognition site of plasminogen binding proteins.
XX The polypeptides are useful in preparing thrombolytic agents for treating
XX blood clotting disorders such as heart attack.
XX
XX Sequence 245 AA;
SQ

Query Match 85.0%; Score 85; DB 21; Length 245;
Best Local Similarity 84.2%; Pred. No. 2,7e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 IVNGEAVPHKXKXQVSLQ 19
Db 16 IVNGEAVFGSPKQVSLQ 34
:|||||:|||||

```

XX PR 22-NOV-1994; 94US-0380501.
XX PR 08-FEB-1995; 95US-0385540.
XX PR 07-JUN-1995; 95US-0486820.
XX PA (PHAI-) PHAIRSON MEDICAL INC.
XX XX Lindblom R, Kay J, Franklin RL, De Faire JR;
XX PI WPI; 1999-561004/47.
XX XX
XX DR Treating acne and eczema using a krill-derived multifunctional enzyme -
XX PT Disclosure: Column 21-22; 42pp; English.
XX PS
XX CC This invention describes a novel method for treating acne and eczema
XX CC using a krill-derived multifunctional enzyme (I) which comprises 2 or
XX CC more of the activities of chymotrypsin, trypsin, collagenase, elastase,
XX CC or exopeptidase and is reactive with cell surface receptors such as
XX CC proteins or glycoproteins. The product of the invention have
XX CC antiseborrheic, anti-acne, dermatological and anti-eczema activity.
XX CC (i) removes or inactivates cell surface receptors (proteins and
XX CC glycoproteins) and adhesion molecules such as ICAM-1 (1-6, CD54), ICAM-2,
XX CC VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD44, CD49, CD62L, CD102 and the
XX CC asialo GM1 ceramide. This sequence represents the N-terminal fragment of
XX CC bovine chymotrypsin which is used to describe the method of the
XX CC invention.
XX SQ Sequence 20 AA;
XX
Query Match 82.0%; Score 82; DB 20; Length 20;
Blast Local Similarity 78.9%; Pred. No. 5, 2e-07;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEAVPHYKXWQVSLQ 19
Db 1 IVNGEDAVPGSWPQVSLQ 19

RESULT 8
AAV93935
ID AAV93935 standard; peptide; 20 AA.
XX AC
XX AC AAV93935;
XX XX
DT 03-OCT-2000 (first entry)
XX DT
XX DE
XX KW Transplantation rejection; hydrolase; graft versus host disease;
XX KW cell surface adhesion molecule; immune reaction; inflammation; shock;
XX KW tumour metastasis; autoimmune disease;
XX KW Krill derived multifunctional enzyme.
XX OS Bos sp.
XX OS
XX PN WO200038708-A1.
XX XX
XX PD 06-JUL-2000.
XX XX
XX PF 23-DEC-1999; 99WO-US30818.
XX XX
XX PR 24-DEC-1998; 98US-0114147.
XX XX
XX PA (PHAI-) PHAIRSON MEDICAL INC.
XX XX
XX PI Franklin RL, St Pierre Y;
XX XX WPI; 2000-452301/39.
XX XX
XX PT Preventing or ameliorating transplantation rejection reactions using
XX PT hydrolase enzymes -
XX XX

```

```

XX PR 22-NOV-1994; 94US-0380501.
XX PR 08-FEB-1995; 95US-0385540.
XX PR 07-JUN-1995; 95US-0486820.
XX PA (PHAI-) PHAIRSON MEDICAL INC.
XX XX Lindblom R, Kay J, Franklin RL, De Faire JR;
XX PI WPI; 1999-561004/47.
XX XX
XX DR Treating acne and eczema using a krill-derived multifunctional enzyme -
XX PT Disclosure: Column 21-22; 42pp; English.
XX PS
XX CC This invention describes a novel method for treating acne and eczema
XX CC using a krill-derived multifunctional enzyme (I) which comprises 2 or
XX CC more of the activities of chymotrypsin, trypsin, collagenase, elastase,
XX CC or exopeptidase and is reactive with cell surface receptors such as
XX CC proteins or glycoproteins. The product of the invention have
XX CC antiseborrheic, anti-acne, dermatological and anti-eczema activity.
XX CC (i) removes or inactivates cell surface receptors (proteins and
XX CC glycoproteins) and adhesion molecules such as ICAM-1 (1-6, CD54), ICAM-2,
XX CC VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD44, CD49, CD62L, CD102 and the
XX CC asialo GM1 ceramide. This sequence represents the N-terminal fragment of
XX CC bovine chymotrypsin which is used to describe the method of the
XX CC invention.
XX SQ Sequence 20 AA;
XX
Query Match 82.0%; Score 82; DB 20; Length 20;
Blast Local Similarity 78.9%; Pred. No. 5, 2e-07;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEAVPHYKXWQVSLQ 19
Db 1 IVNGEDAVPGSWPQVSLQ 19

RESULT 8
AAV93935
ID AAV93935 standard; peptide; 20 AA.
XX AC
XX AC AAV93935;
XX XX
DT 03-OCT-2000 (first entry)
XX DT
XX DE
XX KW Transplantation rejection; hydrolase; graft versus host disease;
XX KW cell surface adhesion molecule; immune reaction; inflammation; shock;
XX KW tumour metastasis; autoimmune disease;
XX KW Krill derived multifunctional enzyme.
XX OS Bos sp.
XX OS
XX PN WO200038708-A1.
XX XX
XX PD 06-JUL-2000.
XX XX
XX PF 23-DEC-1999; 99WO-US30818.
XX XX
XX PR 24-DEC-1998; 98US-0114147.
XX XX
XX PA (PHAI-) PHAIRSON MEDICAL INC.
XX XX
XX PI Franklin RL, St Pierre Y;
XX XX WPI; 2000-452301/39.
XX XX
XX PT Preventing or ameliorating transplantation rejection reactions using
XX PT hydrolase enzymes -
XX XX

```

PS Disclosure; Page 26; 66pp; English.

XX The specification describes a method for preventing or ameliorating
CC transplantation rejection reactions for transplantation of immune cells
CC with other tissues. The method comprises treating a source of immune cells
CC with a protease or hydrolase mixture and administering the treated
CC immune cells to a recipient animal. The hydrolase especially has a
CC preference for inactivating, inactivating or disabling at least
CC one of CD4, CD8, CD25, CD28, CD34, CD45, CD52, an integrin, CD154,
CC CD40 and CD80 in contrast to removing, destroying, inactivating or
CC disabling Tcr. The methods are useful for treating, preventing or
CC disease by using hydrolase enzymes to remove the cell surface adhesion
CC molecules which are involved in triggering the immune reaction in
CC in the diseases. The methods are used for treating or preventing
CC cell-cell or cell-virus adhesion syndrome comprising inflammation, shock,
CC tumour metastases, autoimmune disease, transplantation rejection,
CC reactions or microbial infections. The present sequence represents the
CC N-terminal of a hydrolase, which may be used in the method of the
CC invention.

XX Sequence 20 AA;

Query Match 82.0%; Score 82; DB 21; Length 20;
Best Local Similarity 78.9%; Pred. No. 5.2e-07;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXXMQVSLQ 19
Db 1 IVNGEDAVGSPMQVSLQ 19
|||||:|||||

RESULT 9

AAE07941
D AAE07941 standard; peptide; 20 AA.

AC AAE07941;

DT 01-NOV-2001 (first entry)

DE N-terminal of bovine serine multifunctional hydrolase.

XX Multifunctional hydrolase; rejection reaction; extra-corporeal;
KW therapy; graft versus host disease; transplantation rejection;
KW autoimmune disease; microbial infection; immune disease;
KW cystic fibrosis; chronic obstructive pulmonary disease; COPD;
KW atherosclerosis; cancer; asthma; septic shock; toxic shock syndrome;
KW conjunctivitis; reperfusion injury; pain; immunosuppressive;
KW antibacterial; vasotropic; bovine.

XX Bos sp.

XX US6232088-B1.

XX 15-MAY-2001.

XX 24-DEC-1998; 98US-0220731.

XX 08-FEB-1995; 95US-0385540.

XX 07-JUN-1995; 95US-0486920.

XX 08-FEB-1996; 96US-0600273.

PA (PHAI-) PHAIRSON MEDICAL INC.

XX Franklin RL, St Pierre Y;

XX WPI; 2001-450051/48.

XX Ameliorating, preventing or treating immune rejection reactions, such
PT as graft versus host disease, autoimmune disease, asthma, cancer, by
PT extra-corporeally treating donor tissue with hydrolase such as a
PT protease -

XX Disclosure; Column 15-16; 27pp; English.

XX

CC The present invention relates to a method for ameliorating
CC transplantation rejection reaction. The method comprises treating
CC a rejection ready donor tissue or donor source of immune cells with
CC a hydrolase mixture and administering the treated immune cells to a
CC recipient animal. The hydrolase especially has a preference for
CC rejection reactions such as graft, treating or preventing immune
CC rejection reactions such as graft, treating or preventing immune
CC microbial infection, autoimmune disease, transplantation rejection,
CC obstructive pulmonary disease (COPD), atherosclerosis, cancer, asthma,
CC septic shock, toxic shock syndrome, conjunctivitis, reperfusion injury,
CC and pain in humans. The present sequence is the N-terminal peptide
CC exemplification of the invention.

XX Sequence 20 AA;

Query Match 82.0%; Score 82; DB 22; Length 20;
Best Local Similarity 78.9%; Pred. No. 5.2e-07;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXXMQVSLQ 19
Db 1 IVNGEDAVGSPMQVSLQ 19
|||||:|||||

RESULT 10

AAE54191
D AAE54191 standard; Protein; 146 AA.

AC AAE54191;

DT 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:643.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytotoxic; neuroprotective;
KW neotropic; immunomodulatory; relaxant; contraceptive; gynecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.

XX Homo sapiens.

XX WC200055320-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05989.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-579444/54.

XX N-PSDB; AAC98956.

XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -

XX Claim 11; Page 1081; 1379pp; English.

CC AAC98773 to AAC99211 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigen, given in AAE54018 to
CC AAE54466. The human pancreatic cancer antigen has cytotoxic,
CC neuroprotective, neotropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities and can be used
CC in gene therapy. The polynucleotide and proteins can be used for

CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in an genome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect, and/or target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC92232 to AAC92240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.

XX SQ Sequence 146 AA;

Query Match 82.0%; Score 82; DB 21; Length 146;
 Best Local Similarity 78.9%; Pred. No. 4.9e-06;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 IVNGEEAVPHKXKQVLSQ 19
 |||||:|||||:|||||
 DB 38 IVNGEDAVGSGMPQVLSQ 56

RESULT 11
 AAB98504
 ID AAB98504 standard; Protein; 231 AA.
 XX AAB98504;
 AC AAB98504;
 DT 03-AUG-2001 (first entry)
 XX Human chymotrypsin serine protease catalytic domain.
 DE Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;
 KW tumour antigen-derived gene 15; serine protease; chymotrypsin.
 KW Homo sapiens.
 OS WO200129056-A1.
 PN 26-APR-2001.
 PD 20-OCT-2000; 2000NO-US29095.
 PF 20-OCT-1999; 99US-0421213.
 PR (UYAR-) UNIV ARKANSAS.
 PA O'Brien JV, Tanimoto H;
 XI WPI; 2001-381031/40.
 XX Novel extracellular serine protease, termed tumor antigen-derived gene
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for
 PT diagnosis, treatment, prevention of cancer, particularly breast,
 PT ovarian cancer -
 XX Example 10; Fig 1; 130pp; English.

CC The present invention relates to human tumour antigen-derived gene 15
 CC (TADG-15) protein and coding sequence (see AAB23601 and AAB98500).
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of
 CC 9-20 residues that lack TADG-15 protease activity are useful for
 CC vaccinating an individual against TADG-15, having, suspected of having or
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a
 CC diagnostic or therapeutic target in cancer. The present sequence was used
 CC in a sequence homology alignment with the catalytic domain of TADG-15.

XX SQ Sequence 231 AA;

Query Match 82.0%; Score 82; DB 22; Length 231;
 Best Local Similarity 78.9%; Pred. No. 8.1e-06;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 IVNGEEAVPHKXKQVLSQ 19
 |||||:|||||:|||||
 DB 2 IVNGEDAVGSGMPQVLSQ 20

RESULT 12
 AAU82738
 ID AAU82738 standard; Protein; 263 AA.
 XX AAU82738;
 AC AAU82738;
 DT 23-APR-2002 (first entry)
 XX Amino acid sequence of novel human protease #37.
 DE Human; protease; cancer; immune-related disorder; cardiovascular disease;
 KW neuronal-associated diseases; metabolic disorder; inflammatory disorder;
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;
 KW hypertension; psychiatric disorder; neurological disorder; dyskinesia;
 KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
 KW ocular disease; cytostatic; enzyme.
 XX Homo sapiens.
 OS WO200200860-A2.
 PN 03-JAN-2002.
 PD 26-JUN-2001; 2001WO-US20171.
 PF 26-JUN-2000; 2000US-214047P.
 PR (SUGE-) SUGEN INC.
 PA Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
 XI Charyczak G;
 XX WPI; 2002-139913/18.
 XX N-PSDB; ABR31780.

CC Nucleic acids encoding novel human proteases, useful for useful for
 CC treating diseases and disorders such as cancers, immune-related
 CC diseases and disorders, cardiovascular disease (e.g. restenosis) and
 CC inflammatory disorders -
 XX Claim 6; Fig 2M; 313pp; English.

CC The present invention relates to the isolation of novel human
 CC proteases, and the nucleic acids encoding them. The sequences of
 CC the invention are useful for treating diseases and disorders such as
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
 CC (e.g. inflammatory diseases and diseases), cardiovascular diseases
 CC (e.g. restenosis and coronary atherosclerosis), brain or neuronal-associated
 CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
 CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
 CC peripheral nervous system disorders, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypotension,
 CC hyper-emotion, psychotic disorders, neurological disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
 CC The nucleic acids and polypeptides are also useful for treating viral
 CC infections caused by human immunodeficiency virus (HIV), and non-viral
 CC infections such as ocular disease (e.g. glaucoma) and macular
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
 CC the invention.

XX SQ Sequence 263 AA;

```

Query Match      82.0%; Score 82; DB 23; Length 263;
Best Local Similarity 78.9%; Pred. No. 9.4e-06;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
      |||||
Db 34 IVNGEAVPGSNWQVSLQ 52

RESULT 13
AAP60060
ID AAP60060 standard; Protein; 242 AA.
XX
AC AAP60060;
XX
DT 06-SEP-1991 (first entry)
XX
DE Sequence of human pancreatic elastase IIA.
XX
EN Enzyme; serum lipoprotein metabolism.
XX
OS Homo sapiens.
XX
PN EP198645-A.
XX
PD 22-OCT-1986.
XX
PF 07-APR-1986; 86EP-0302557.
XX
PR 23-OCT-1985; 85JP-0236686.
XX
PR 05-APR-1985; 85JP-0072308.
XX
PR 27-APR-1985; 85JP-0091986.
XX
PR 26-JUL-1985; 85JP-0163964.
XX
PR 02-DEC-1985; 85JP-0271128.
XX
XX (SANY ) SANYO KK.
XX
PI Takiguchi Y, Tani T, Kawashima I, Erukawa H, Ohnine T;
XX Ohsumi J;
DR WPI: 1986-280300/43.
XX N-PSDB; AAN60074.
XX
XX New genetically engineered human pancreatic elastase - obtd.
XX using hosts modified DNA coding for enzyme
XX
XX Claim 2; Page 6-7; 45pp; English.
XX
XX The genetically engineered prod. can eliminate the dependency on
XX human pancreas supplies for the elastase, and avoids antibody
XX CC formation and possibility of anaphylaxis using porcine elastase.
XX
SQ Sequence 242 AA;

Query Match      80.0%; Score 80; DB 7; Length 242;
Best Local Similarity 68.4%; Pred. No. 1.9e-05;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
      :|||:||||:
Db 1 VVHGEDAVPSWQVSLQ 19

RESULT 14
AAB11711
ID AAB11711 standard; Protein; 264 AA.
XX
AC AAB11711;
XX
DT 23-OCT-2000 (first entry)
XX
DE Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.
XX
XX BSSP5; serine protease; human; mBSSP5; mouse; mBSSP5; brain;
XX diagnostic marker; antibody; tumor; animal; Alzheimer's disease;
XX epilepsy; cancer; inflammation; infertility; pancreatitis;
XX prostatic hypertrophy.
XX
OS Mus sp.
XX
XX WO20031243-A1.
XX
XX 02-JUN-2000.
XX
XX 19-NOV-1999; 99WO-JP06473.
XX
XX 20-NOV-1998; 98JP-0347806.
XX (FUSO ) FUSO PHARM IND LTD.
XX
XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
XX WPI: 2000-400058/34.
XX N-PSDB; AAA61734.
XX
XX Serine proteases BSSP5, useful in detecting homologs, mutants and
XX polymorphic variants as markers for diagnosis of e.g. Alzheimer's
XX disease, epilepsy, cancer and inflammation, using blood, urine,
XX pancreas or other tissues.
XX
XX Claim 3; Page 55-56; 70pp; Japanese.
XX
XX The invention relates to novel serine proteases designated BSSP5
XX (AAB1170-AAB1171), and to nucleic acids encoding them (AAA61733-A61734).
XX The invention also relates to vectors and transformants comprising BSSP5
XX nucleotides; and to animals in which the expression level of BSSP5
XX can be varied; and BSSP5 knockout mouse. The invention additionally
XX encompasses anti-BSSP5 antibody, methods of production of such
XX antibodies, methods of BSSP5 detection using the antibodies, and the
XX use of BSSP5 proteins or fragments as diagnostic markers, and the
XX medical conditions, e.g., pancreatitis. A method for detecting
XX pancreatitis comprising measuring BSSP5 concentration in the blood or
XX urine, and a pancreatitis diagnostic agent containing an anti-BSSP5
XX antibody is also disclosed. Nucleotides encoding BSSP5 were initially
XX isolated in a human brain cDNA library using degenerate PCR primers
XX (AAA61744-A61745) based on conserved regions of serine proteases. The
XX BSSP5 serine proteases and nucleotides encoding them are useful in
XX detecting homologs, mutants and polymorphic variants in biological
XX samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
XX pancreas and spleen) as diagnostic markers for conditions such as
XX pancreatic disease, epilepsy, cancer, inflammation, infertility.
XX The invention also encompasses a prostatic hypertrophy. Sequence AAB11710 represents
XX human BSSP5 (mBSSP5), and sequence AAB11711 represents murine BSSP5
XX (mBSSP5).
XX
SQ Sequence 264 AA;

Query Match      80.0%; Score 80; DB 21; Length 264;
Best Local Similarity 78.9%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
      |||||
Db 34 IVNGEAVPGSNWQVSLQ 52

RESULT 15
AAP70759
ID AAP70759 standard; Protein; 270 AA.
XX
AC AAP70759;
XX
DT 19-APR-1991 (first entry)
XX
DE Human pancreas elastase-1.

```

XX	Lipoprotein metabolism; lipid.
XX	
OS	Sus scrofa.
OS	
PH	Key Location/Qualifiers
FT	Misc-difference 2..28
PT	/note= "May be absent"
XX	
PX	JP6200276-A.
PN	
PD	06-JAN-1987.
XX	
PF	25-JUN-1985; 85JP-0138494.
XX	
PR	25-JUN-1985; 85JP-0138494.
XX	(SANY) SANKYO KK.
PA	
DR	WPI; 1987-040875/06.
XX	N-PSDB; AAN71122.
XX	
PPT	Pancreas elastase prodn. used to improve lipoprotein metabolism -
PT	comprises isolating RNA coding elastase, synthesising single and
PT	double chain cDNA and introducing recombinant into host
XX	
PS	Disclosure; Page 481; 18pp; Japanese.
XX	
CC	Sequence encoding pancreas elastase may be used to transform an
CC	expression system. The product is useful in rejuvenating the
CC	elasticity of the arterial wall, and improving unusual serum lipid
CC	levels and lipoprotein metabolism.
XX	
SQ	Sequence 270 AA;
Query Match	80.0%; Score 80; DB 8; Length 270;
Best Local Similarity	68.4%; Pred. NO. 2.1e-05;
Matches 13; Conservative	4; Mismatches 2; Indels 0; Gaps 0;
Oy	1 IVNGEAVPHKXMKVQS LQ 19 : :
Dd	29 VHGEDAVPXSMPQVSLQ 47
Search completed:	February 12, 2003, 10:22:24
Job time :	14.4328 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 5.01493 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-4

Perfect score: 100

Sequence: 1 IVNGEEAVPHXKXQVSLQ 20

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	97.0	263	2 S47537	chymotrypsin (EC 3
2	95	95.0	244	2 S72219	chymotrypsin B - A
3	85	85.0	31	2 B33257	elastase (EC 3.4
4	85	85.0	245	1 KYBOA	chymotrypsin (EC 3
5	85	85.0	270	2 B29934	pancreatic elastase
6	83	83.0	23	2 P00036	serine proteinase
7	83	83.0	28	2 A61529	chymotrypsin (EC 3
8	82	82.0	245	1 KYBOB	chymotrypsin (EC 3
9	82	82.0	263	2 A11195	chymotrypsin (EC 3
10	82	82.0	263	2 A11299	chymotrypsin (EC 3
11	81	81.0	263	1 KYRFB	chymotrypsin (EC 3
12	81	81.0	260	2 B61333	pancreatic elastase
13	80	80.0	258	2 Q01473	pancreatic elastase
14	80	80.0	258	2 S68934	pancreatic elastase
15	77	77.0	268	2 S68926	pancreatic elastase
16	77	77.0	268	2 S68926	pancreatic elastase
17	75	75.0	240	1 CPBOA3	pancreatic elastase
18	74	74.0	269	2 A25528	pancreatic elastase
19	74	74.0	271	2 A25528	pancreatic elastase
20	73	73.0	126	2 A23473	pancreatic elastase
21	72	72.0	271	1 ELR72	chymotrypsin-like
22	71	71.0	269	2 A26823	chymotrypsin-like
23	70	70.0	264	2 I38136	chymotrypsin-like
24	69	69.0	269	2 C26823	chymotrypsin-like
25	67	67.0	1004	2 T30338	pancreatic elastase
26	65	65.0	20	2 S33787	pancreatic elastase
27	65	65.0	25	2 S12997	pancreatic elastase
28	65	65.0	1524	2 T30337	polyprotein - Afri
29	64	64.0	460	2 B61545	plasmin (EC 3.4.21

plasmin (EC 3.4.21)
chymotrypsin (EC 3
plasmin (EC 3.4.21
apolipoprotein(a)
apoprotein(a) (EC
pancreatic elastase
chymotrypsin (EC 3
plasmin (EC 3.4.21
plasmin (EC 3.4.21
plasmin (EC 3.4.21
trypsin (EC 3.4.21
mast cell tryptase
plasmin (EC 3.4.21
trypsin (EC 3.4.21
trypsin (EC 3.4.21
acrosin (EC 3.4.21
plasmin (EC 3.4.21

ALIGNMENTS

RESULT 1
S47537
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
C:Species: Gadus morhua (Atlantic cod)
C>Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
R:Accession: S47537; S43163
R:Gudmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.
Biochim. Biophys. Acta 1219, 211-214, 1994
A>Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
A:Reference number: S47537; MUID:94368860; PMID:8086467
A:Accession: S47537
A:Molecule type: mRNA
A:Residues: 1-263 <GDS>
A:Cysteine sites: SMDLX78490; NID:9468750; PID:CAA55242.1; PID:9468751
C:Superfamily: trypsin, trypsin homology
C:Keywords: hydrolyse, proteinase, serine proteinase
F:1-18/Domain: signal, sequence #status predicted <SD>
F:19-263/Product: chymotrypsin #status predicted <SD>
F:34-256/Domain: trypsin homology <TRY>
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 97.0%; Score 97; DB 2; Length 263;
Best Local Similarity 89.5%; Pred. No. 1.2e-08;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXKXQVSLQ 19
Db 34 IVNGEEAVPHXKXQVSLQ 52

RESULT 2

S72219

chymotrypsin B - Atlantic cod (fragments)

C:Species: Gadus morhua (Atlantic cod)

C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Aug-1998

R:Accession: S72219

R:Lech-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Holrup, P.

Biochim. Biophys. Acta 1297, 49-56, 1996

A>Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.

A:Reference number: S72219; MUID:96439045; PMID:8841380

A:Accession: S72219

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-14;15-244 <LET>

A:Superfamily: trypsin; trypsin homology

F:15-237/Domain: trypsin homology <TRY>

Query Watch 95.0%; Score 95; DB 2; Length 244;

Best Local Similarity 89.5%; Pred. No. 2.4e-08;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXKXQVSLQ 19

```

Db      15 IVNGEEAVPHXWQVMSLQ 33
|||||
RESULT 3
33257
pancreatic elastase (EC 3.4.21.36) III - pig (fragment)
N/Alternate names: proteinase E
C/Comment: Sue scrota domestica (domestic pig)
C/Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 30-Jun-1995
C/Accession: B33257
R/Aviles, F.X.; Pascual, R.; Salva, M.; Bonicel, J.; Puigserver, A.
Biochem. Biophys. Res. Commun. 163, 1191-1196, 1989
A/Title: Generation of a subunit III-like protein by autolysis of human and porcine pro-
A/Reference: B33257, PMID:89392022, PMID:2675835
A/Accession: B33257
A/Keywords: chymotrypsin; trypsin; trypsin homology
A/Molecule type: protein
A/Residues: 1-31 <AVI>
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase

Query Match      85.0%; Score 85; DB 2; Length 31;
Best Local Similarity 73.7%; Pred. No. 1.2e-07;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY      1 IVNGEEAVPHXWQVMSLQ 19
|||||
Db      12 VNGEDAVPSWQVMSLQ 30
|||||

RESULT 4
KYBOA
chymotrypsin (EC 3.4.21.1) A precursor - bovine
N/Alternate names: chymotrypsinogen A
C/Species: Bos primigenius taurus (cattle)
C/Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999
C/Accession: A90235; A93158; S29650; A00952
R/Brown, J.R.; Hartley, B.S.
Biochem. J. 101, 214-228, 1966
A/Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphid
A/Reference number: A90235; PMID:67181721; PMID:5971783
A/Accession: A90235
A/Molecule type: protein
A/Residues: 1-101; N', 103-245 <BR>
R/Blow, D.M.; Birkoft, J.J.; Hartley, B.S.
Nature 211, 337-340, 1969
A/Title: Role of a buried acid group in the mechanism of action of chymotrypsin.
A/Reference number: A93158; PMID:5910266; PMID:5764436
A/Accession: A93158; revision to residue 102
R/Meloun, B.; Kludis, R.; Kettner, V.; Moravik, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Som
Biochim. Biophys. Acta 130, 543-546, 1966
A/Title: Covariant structure of bovine chymotrypsinogen A.
A/Reference number: A90572; PMID:67183948; PMID:5972866
A/Accession: A93158
A/Molecule type: protein
A/Residues: 1-101; N', 103-245 <WEL>
A/Note: disulfide bonds were determined
R/Curruzola, F.; Accensi, P.; Barra, D.; Bolognesi, M.; Menegatti, E.; Sarti, P.; Schne
Biochim. Biophys. Acta 1161, 201-208, 1993
A/Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catalyti
A/Reference number: S29650; PMID:93160238; PMID:8431470
A/Accession: S29650
A/Molecule type: protein
A/Residues: 1-12; 16-27; 149-160; 181-200 <CUT>
R/Smillie, L.B.; Hartley, B.S.
Biochem. J. 101, 232-241, 1966
A/Title: Histidine sequences in the active centres of some 'serine' proteinases.
A/Reference number: A90236; PMID:67181723; PMID:5971785
A/Accession: A90236
A/Contents: annotation; active site
R/Birkoft, J.J.; Blow, D.M.; Henderson, R.; Steitz, T.A.
Philos. Trans. R. Soc. Lond. B257, 67-76, 1970
A/Title: The structure of alpha-chymotrypsin.

```

```

A/Reference number: A93754
A/Contents: annotation; X-ray crystallography
C/Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.
C/Comment: Trypsin cleavage after Arg-15 results in a fully active enzyme (pi-chymotrypsin)
delta-chymotrypsin; further chymotryptic cleavage liberates the dipeptide Thr-147 and Asn-148
directly from chymotrypsinogen, which leads to the degraded form neochymotrypsin
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F/1-13; 16-146; 149-245/Product: alpha-chymotrypsin #status experimental <ZYM>
F/16-238/Domain: trypsin homology <TRY>
F/1-122; 45-58; 136-201; 168-182; 191-220/Disulfide bonds: #status experimental
F/57; 102; 195/Active site: His, Asp, Ser #status experimental

Query Match      85.0%; Score 85; DB 1; Length 245;
Best Local Similarity 84.2%; Pred. No. 1.1e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1 IVNGEEAVPHXWQVMSLQ 19
|||||
Db      16 IVNGEEAVPGSWQVMSLQ 34
|||||

RESULT 5
B29934
pancreatic elastase (EC 3.4.21.36) IIIB precursor - human
N/Alternate names: 35K glycoprotein; pancreatic protein P35; protein G32; proteinase E
C/Species: Homo sapiens (man)
C/Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 22-Jun-1999
C/Accession: B29934; A93316; A57206; A33257; A28932; S04399; S04490
R/Fani, T.; Ohnuma, T.; Takiguchi, Y.
J. Biol. Chem. 263, 1231-1239, 1988
A/Title: Identification of a novel class of elastase isozyme, human pancreatic elastase I
A/Reference number: A92664; PMID:88087253; PMID:2826474
A/Accession: B29934
A/Molecule type: mRNA
A/Residues: 1-270 <TAN>
A/Cross-references: GB:M18692; NID:G607029; PIDN:AAA58454.1; PID:gl82035
R/Shen, W.; Fletcher, T.S.; Laxman, C.
Biochemistry 26, 3447-3452, 1987
A/Title: Primary structure of human pancreatic protease E determined by sequence analysis
A/Reference number: A90516; PMID:8800545; PMID:3477287
A/Accession: A90516
A/Molecule type: mRNA
A/Residues: 'G', '5-63', 'G', '65-78', 'W', '80-118', 'G', '120-163', 'P', '165-270 <SHE>
R/Fletcher, T.S.
submitted to GenBank, August 1987
A/Reference number: A94507
A/Contents: revision to residue 119
A/Accession: A27206
A/Molecule type: DNA
A/Residues: 'G', '5-63', 'G', '65-78', 'W', '80-163', 'P', '165-270 <FILE>
R/Aviles, F.X.; Pascual, R.; Salva, M.; Bonicel, J.; Puigserver, A.
Biochem. Biophys. Res. Commun. 163, 1191-1196, 1989
A/Title: Generation of a subunit III-like protein by autolysis of human and porcine prope
A/Reference number: A33257; PMID:89392022; PMID:2675835
A/Accession: A33257
A/Molecule type: protein
A/Residues: 18-57 <AVI>
R/Guy-Crotte, O.; Barthe, C.; Basso, D.; Fournet, B.; Figarella, C.
Biochem. Biophys. Res. Commun. 156, 318-322, 1988
A/Title: Characterization of two glycoproteins of human pancreatic juice: P35, a truncate
A/Reference number: A28932; PMID:89045862; PMID:3178837
A/Accession: A28932
A/Molecule type: protein
A/Residues: 'X', '34-52', 'X', '54-55', 'XX', '59-63 <GVY>
R/Moulard, M.; Kerfelec, P.
FEBS Lett. 250, 166-170, 1989
A/Title: Identification of procarboxypeptidase A-truncated protease E binary complex
A/Reference number: S04999; PMID:89355560; PMID:2755124
A/Accession: S04999
A/Molecule type: protein
A/Residues: 31-50 <MOU>

```


R.Wendorf, P.; Geyer, R.; Szlegoleit, A.; Linder, D.
 FBS Lett. 249, 275-278, 1989
 A>Title: Localization and characterization of the glycosylation site of human pancreatic trypsinogen
 A:Reference number: S04490; MUID:89289996; PMID:12737288
 A:Accession: S04490
 A:Molecule type: Protein
 A:Residues: 1-261
 A:Superfamily: trypsin homology
 C:Keywords: glycoprotein; serine proteinase; zymogen
 F:1-17/Domain: signal sequence #status predicted <ACT>
 F:18-28/Domain: activation peptide #status predicted <ACT>
 F:29-270/Product: (or 31-270) pancreatic elastase IIIB #status predicted <MAT>
 F:29-263/Domain: trypsin homology <TRY>
 F:73,123,217/Active site: His, Asp, Ser #status predicted
 F:114/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:153/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 85.0%; Score 85; DB 2; Length 270;
 Best Local Similarity 73.7%; Pred. No. 1.2e-06;
 Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IUNGEEAVPHKXKQVSLQ 19
 DB 29 IUNGEEAVPHKXKQVSLQ 47

RESULT 6
 PU0036
 serine proteinase (EC 3.4.21.-) - bovine (fragment)
 N:Alternate names: trypsin-like proteinase bPTLP
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 22-Apr-1995
 C:Accession: PU0036; PU0039
 R:Teuli, A.; Sakiyama, K.; Bdzawa, K.; Nagata, K.; Sasaki, Y.; Nagamune, H.; Matsuda, Y.
 submitted to JPTD, September 1994
 A:Description: Purification and characterization of a novel serine proteinase from bovine pancreas
 A:Reference number: PU0036
 A:Accession: PU0036
 A:Residues: 1-253
 A:Molecule type: Protein
 A:Superfamily: trypsin homology
 C:Keywords: hydrolase; serine proteinase

Query Match 83.0%; Score 83; DB 2; Length 23;
 Best Local Similarity 63.2%; Pred. No. 1.8e-07;
 Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IUNGEEAVPHKXKQVSLQ 19
 DB 1 VNGEDAVPGSNPQVSLQ 19

RESULT 7
 A61520
 chymotrypsin (EC 3.4.21.1) - Atlantic cod (fragments)
 C:Species: Gadus morhua (Atlantic cod)
 C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 11-May-2000
 C:Accession: A61529
 R:Aggairsson, B.; Bjarnason, J.B.
 Comp. Biochem. Physiol. B 99, 327-335, 1991
 A>Title: Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua)
 A:Reference number: A61529; MUID:92111252; PMID:1764912
 A:Accession: A61529
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-28 <ASG>
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 83.0%; Score 83; DB 2; Length 28;
 Best Local Similarity 87.5%; Pred. No. 2.2e-07;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IUNGEEAVPHKXKQV 16
 DB 13 IUNGEEAVPHKXKQV 28

RESULT 8

KYBOB

chymotrypsin (EC 3.4.21.1) B precursor - bovine

N:Alternate names: chymotrypsinogen B

C:Species: Bos primigenius taurus (cattle)

C:Date: 08-Oct-1991 #sequence_revision 08-Oct-1991 #text_change 18-Jul-1997

C:Accession: A00953

R:Smith, L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.

N:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen

A:Reference number: A00953; MUID:68238908; PMID:5649671

A:Accession: A00953

A:Molecule type: protein

A:Residues: 1-245 <SM>

C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acini

C:Comment: The first activation cleavage, leading to pi-chymotrypsin B, occurs in the ear

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-15/Domain: propeptide #status experimental <PRO>

F:16-245/Product: chymotrypsin B #status experimental <MAT>

F:16-238/Domain: trypsin homology <TRY>

F:12,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental

F:57,102,195/Active site: His, Asp, Ser #status experimental

C:Accession: A31299
 R:Ontita, N., Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsub
 Biochem. Biophys. Res. Commun. 158, 569-575, 1989
 A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinog
 A:Reference number: A31299, MUID:89134264, PMID:2917002
 A:Accession: A31299
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-263 <TOM>
 A:Cross-references: GB:M24400, NID:g181189; PIDN:AAAS2128.1; PID:g181190
 C:Genetics:
 A:Gene: GDB:CTRB1, CTRB
 A:Cross-references: GDB:119820; OMIM:118890
 A:Map position: 16q23.1-16q23.1
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 82.0%; Score 82; DB 2; Length 263;
 Best Local Similarity 78.9%; Pred. No. 3,8e-06;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKQVSLQ 19
 |||||:|||||
 Db 34 IVNGEDAVPGSNPQVSLQ 52

RESULT 11
 KYRTB
 chymotrypsin (EC 3.4.21.1) B precursor - rat
 N:Alternate names: chymotrypsinogen B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
 C:Accession: A22658
 R:Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.
 J. Biol. Chem. 259, 14265-14270, 1984
 A:Title: Isolation and sequence of a rat chymotrypsin B gene.
 A:Reference number: A22658; MUID:85054881; PMID:6209274
 A:Accession: A22658
 A:Molecule type: DNA
 A:Residues: 1-263 <BEL>
 A:Cross-references: GB:K02298, NID:g203653; PIDN:AAA98732.1; PID:g203654
 C:Genetics:
 A:Introns: 18/1, 52/3; 79/2, 105/3, 166/1; 210/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
 F:10-31/Domain: signal peptide #status predicted <SIG>
 F:19-31/Domain: signal peptide #status predicted <PRO>
 F:34-263/Product: chymotrypsin B #status predicted <MAT>
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 81.0%; Score 81; DB 1; Length 263;
 Best Local Similarity 73.7%; Pred. No. 5,7e-06;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKQVSLQ 19
 |||||:|||||
 Db 34 IVNGEDAVPGSNPQVSLQ 52

RESULT 12
 B61333
 chymotrypsin (EC 3.4.21.1) - edible frog (fragment)
 C:Species: Rana esculenta (edible frog)
 C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
 C:Accession: B61333
 R:Pies, W.; Zwilling, R.; Woodbury, R.G.; Neurath, H.
 FEBS Lett. 109, 45-49, 1980
 A:Title: Amino-terminal amino acid sequences and the evolution of frog (Rana esculenta)
 A:Reference number: A61333; MUID:80113255; PMID:6965480
 A:Accession: B61333

A:Molecule type: protein
 A:Residues: 1-20 <PIE>
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 80.0%; Score 80; DB 2; Length 20;
 Best Local Similarity 78.9%; Pred. No. 4,9e-07;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKQVSLQ 19
 |||||:|||||
 Db 1 IVNGENAVPGSNPQVSLQ 19

RESULT 13
 JQ1473
 pancreatic elastase (EC 3.4.21.36) IV precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 22-Jun-1999
 C:Accession: JQ1473; S23783
 R:Xang, J.; Wiegand, U.; Mueller-Hill, B.
 Gene 110, 181-187, 1992
 A:Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.
 A:Reference number: JQ1471; MUID:92165057; PMID:1537555
 A:Accession: JQ1473
 A:Molecule type: mRNA
 A:Residues: 1-268 <KAN>
 A:Cross-references: EMBL:X59014, NID:g56090; PIDN:CAA11753.1; PID:g56091
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase; zymogen
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-29/Domain: activation peptide #status predicted <ACT>
 F:30-268/Product: pancreatic elastase IV #status predicted <MAT>
 F:34-262/Domain: trypsin homology <TRY>
 F:74,121,216/Active site: His, Asp, Ser #status predicted

Query Match 80.0%; Score 80; DB 2; Length 268;
 Best Local Similarity 68.4%; Pred. No. 8,5e-06;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKQVSLQ 19
 |||||:|||||
 Db 30 VVGEDAVPGSNKQVSLQ 48

RESULT 14
 A29934
 pancreatic elastase (EC 3.4.21.36) IIIA precursor - human
 N:Alternate names: protease B
 C:Species: Homo sapiens (man)
 C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Jun-2000
 C:Accession: A29934; JX0045
 R:Tani, T.; Ohsumi, J.; Mita, K.; Takiguchi, Y.
 J. Biol. Chem. 263, 1231-1239, 1988
 A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase I
 A:Reference number: A29664; MUID:88087253; PMID:2826474
 A:Accession: A29934
 A:Molecule type: DNA
 A:Residues: 1-270 <TAN>
 A:Cross-references: GB:J03516
 R:Shirasu, Y.; Takemura, K.; Yoshida, H.; Sato, Y.; Iijima, H.; Shimada, Y.; Mikayama, T.
 J. Biochem. 104, 259-264, 1988
 A:Title: Molecular cloning of complementary DNA encoding one of the human pancreatic prot
 A:Reference number: JX0045; MUID:89034017; PMID:2460440
 A:Accession: JX0045
 A:Molecule type: mRNA
 A:Residues: 1-240, 'G', 242-270 <SHI>
 A:Cross-references: GB:D00306, NID:g220013; PIDN:BA00212.1; PID:g220014
 C:Comment: This enzyme is an alanine-specific serine proteinase that has little elastolyt
 C:Genetics:
 A:Introns: 15/1; 43/3; 76/2; 121/2; 167/1; 214/3; 265/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; serine proteinase

F1-15/Domain: signal sequence #status predicted <SIG>
 F16-28/Domain: propeptide #status predicted <PRO>
 F29-270/Product: pancreatic elastase IIA #status experimental <NAT>
 F29-263/Domain: trypsin homology <TRY>
 F75,123,217/Active site: His, Asp, Ser #status predicted

Query Match 80.0%; Score 80; DB 2; Length 270;
 Best Local Similarity 69.4%; Pred. No. 8,6e-06;
 Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVNGEAVPHKWKQVSLQ 19
 :||:||||:|||||
 Db 29 VVHGDAVPSPKQVSLQ 47

RESULT 15

S68826
 pancreatic elastase (EC 3.4.21.36) isoform 2 precursor - human
 A:Alternate names: caldesin isoform 2
 C:Species: Homo sapiens (man)
 C:Accession: S68826
 R:Tomomura, A.; Kiyama, M.; Itoh, H.; Yoshino, I.; Tomomura, M.; Nishii, Y.; Noikura, T.
 FEBS Lett. 366: 26-28, 1995
 A:Title: Molecular cloning and expression of human caldesin.
 A:Reference number: S68825; MUID:96221265; PMID:8635596
 A:Accession: S68826
 A:Molecule type: mRNA
 A:Residues: 1-268 <TOM>
 A:Experimental source: pancreas
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; serine proteinase; zymogen
 F1-16/Domain: signal sequence #status predicted <SIG>
 F17-29/Domain: propeptide #status predicted <PRO>
 F30-266/Product: pancreatic elastase isoform 2 #status predicted <NAT>
 F30-262/Domain: trypsin homology <TRY>

Query Match 77.0%; Score 77; DB 2; Length 268;
 Best Local Similarity 63.2%; Pred. No. 2.7e-05;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVNGEAVPHKWKQVSLQ 19
 :||:||||:|||||
 Db 30 VVGGDARPHSPKQVSLQ 48

Search completed: February 12, 2003, 10:28:49
 Job time : 5.01493 secs

THIS PAGE BLANK (USP 1)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:04:45 ; Search time 2.68657 Seconds
(without alignment)
308.768 Million cell updates/sec

Title: US-10-036-371-4

Perfect score: 100

Sequence: 1 IVNCEAAPHXMYQVSLQX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	97.0	263	1 CTRA_GADMO	P47796 gadus morhu
2	95	95.0	245	1 CTRB_GADMO	P80646 gadus morhu
3	97	87.0	253	1 CAC3_BOVIN	P05805 bos taurus
4	85	85.0	245	1 CTRA_BOVIN	P00766 bos taurus
5	85	85.0	270	1 ELB3_HUMAN	P08861 homo sapien
6	82	82.0	245	1 CTRB_BOVIN	P00767 bos taurus
7	82	82.0	263	1 CTR2_CANFA	P04813 canis famil
8	81	82.0	263	1 CTRB_HUMAN	P17538 homo sapien
9	80	82.0	263	1 CTRA_RAT	P07338 rattus norv
10	80	80.0	263	1 ELB2_HUMAN	P09093 homo sapien
11	78	78.0	268	1 CTRB_HUMAN	P55091 rattus norv
12	77	77.0	268	1 CTRB_HUMAN	P55091 rattus norv
13	74	74.0	269	1 ELB2_HUMAN	P08217 homo sapien
14	74	74.0	271	1 EL2_RAT	P02408 mus musculu
15	72	72.0	271	1 EL2_RAT	P02408 mus musculu
16	71	71.0	269	1 EL2_BOVIN	P02408 mus musculu
17	71	71.0	269	1 EL2_PIG	P08419 sus scrofa
18	70	70.0	269	1 CTRB_HUMAN	P03113 homo sapien
19	69	69.0	269	1 ELB2_HUMAN	P03113 homo sapien
20	66	66.0	492	1 TMS2_HUMAN	O15393 homo sapien
21	65	65.0	201	1 ELAS_GADMO	P32197 gadus morhu
22	65	65.0	275	1 TRV1_PIG	P32197 gadus morhu
23	63	63.0	343	1 PLAN_SHEEP	Q9n2d1 sus scrofa
24	63	63.0	343	1 PLAN_SHEEP	P06667 sus scrofa
25	62	62.0	437	1 COG1_CHLOP	P34153 chionocete
26	62	62.0	437	1 COG1_CHLOP	P34153 chionocete
27	62	62.0	810	1 PLAN_MOUSE	P08419 sus scrofa
28	62	62.0	1420	1 APOA_HUMAN	P12545 macaca mula
29	62	62.0	4548	1 APOA_HUMAN	P14417 macaca mula
30	61	61.0	410	1 PLAN_HUMAN	P08519 homo sapien
31	60	60.0	333	1 PLAN_CANFA	P08068 canis famil
32	60	60.0	812	1 PLAN_BOVIN	P08068 canis famil
33	60	60.0	812	1 PLAN_MOUSE	P20918 mus musculu

34	59	59.0	270	1 TRV1_MERUN	P50342 meriones un
35	59	59.0	273	1 MCT7_RAT	P27435 rattus norv
36	59	59.0	338	1 PLAN_HORSE	P80010 equus caball
37	58	58.0	273	1 MCT7_MOUSE	O02841 mus musculu
38	58	58.0	273	1 PLAN_ERIEU	O02841 mus musculu
39	57	57.0	235	1 TRV1_HUMAN	O02841 mus musculu
40	57	57.0	267	1 ELNE_HUMAN	P08246 homo sapien
41	57	57.0	275	1 TRB1_HUMAN	P02311 homo sapien
42	57	57.0	275	1 TRB2_HUMAN	P02311 homo sapien
43	57	57.0	275	1 TRV1_HUMAN	P15157 homo sapien
44	56	56.0	254	1 KLK4_HUMAN	Q9Y5K2 homo sapien
45	56	56.0	274	1 MCT6_RAT	P50343 rattus norv

ALIGNMENTS

RESULT 1

ID	CTRA_GADMO	STANDARD	PRT	263 AA.
AC	P47796;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DB	Chymotrypsin A precursor (EC 3.4.21.1)			
OS	Gadus morhua (Atlantic cod)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Necnomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	[1]_TaxID=8049;			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Pyloric caeca;			
RX	GMDLINE=94368860; PubMed=8086467;			
RA	Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,			
RA	"Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";			
RL	Biochim. Biophys. Acta 1219:211-214(1994).			
RN	[2]			
RP	SEQUENCE OF 19-30 AND 34-49.			
RC	TISSUE=Pyloric caeca;			
RX	GMDLINE=92111252; PubMed=1764912;			
RA	Aggeston B., Bjarnason J.B.;			
RA	Structural and kinetic properties of chymotrypsin from Atlantic cod			
RL	(Gadus morhua). Comparison with bovine chymotrypsin.";			
RL	Canis familiaris. Physiol. 998:327-335(1991).			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr -Xaa, Trp -Xaa,			
CC	Phe -Xaa, Leu -Xaa; Extracellular			
CC	-1- SUBCELLULAR LOCATION: Extracellular			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL: X78490; CAA5242.1; -			
DR	RSSP; P00766; 1 CHG.			
DR	MERCOPS; S01152.			
DR	InterPro; IPR001314; Chymotrypsin			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin_1			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_Spc; 1			
DR	PROSITE; PS00240; TRYPsin DOM; 1			
DR	PROSITE; PS00134; TRYPsin HIS; 1			
DR	PROSITE; PS00135; TRYPsin SER; 1			
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.			
FT	SIGNAL			
FT	CHAIN	1	263	CHYMOTRYPSIN A.

FT	ACT_SITE	195	195		CHARGE RELAY SYSTEM (BY SIMILARITY) .
FT	DISULFID	1	121		BY SIMILARITY.
FT	DISULFID	42	58		BY SIMILARITY.
FT	DISULFID	135	151		BY SIMILARITY.
FT	DISULFID	167	182		BY SIMILARITY.
FT	DISULFID	191	220		BY SIMILARITY.
FT	CONFLICT	9	11		QVT -> VIS (IN REF. 2) .
FT	CONFLICT	26	26		S -> T (IN REF. 2) .
FT	CONFLICT	28	29		PW -> Y (IN REF. 2) .
FT	SEQUENCE	245 AA;	26260 WM;	74PEOD425517AB02 CRC64;	
Query Match		95.0%	Score 95;	DB 1;	Length 245;
Best Local Similarity		89.5%;	Pred. No. 5.2e-08;		
Matches 17;	Conservative	0;	Mismatches 2;	Indels	0; Gaps 0;
Oy	1	I VNGEEAVPHXKXVSLQ	19		
Dd	16	I VNGEEAVPHSWPQVSLQ	34		
RESULT 3					
CAC3_BOVIN					
ID _CAC3_BOVIN	STANDARD;	PRT;	253 AA.		
IC AC P05805;					
DT 01-NOV-1988 (Rel. 09, Created)					
DT 15-DEC-1998 (Rel. 37, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE PROTEINASE B precursor (Procarboxypeptidase A complex component					
DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).					
DE Bos taurus (Bovine).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;					
OC Bovidae; Bovinae; Bos.					
OX NCBI_TaxId=9913;					
OX [1]					
RN SEQUENCE OF 1-25:					
RX MEDLINE=9109520; PubMed=2269366;					
RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,					
RA "Autoglyser A."					
RT Polysaccharide of proproteinase B in bovine procarboxypeptidase A ternary					
RL complex gives rise to subunit III."					
RL [1] (JBS Lett. 27:37-41(1990).					
RN SEQUENCE OF 14-253 AND DISULFIDE BONDS.					
RX MEDLINE=86220198; PubMed=3519215;					
RA Verot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G.;					
RT "Amino acid sequence and disulfide bridges of subunit III, a					
RL defective endopeptidase present in the bovine pancreatic 6 S					
RL procarboxypeptidase A complex."					
EX J. Biochem. 157:91-99(1986).					
[3]					
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) .					
RX MEDLINE=94222022; PubMed=8168476;					
RA Pignol D., Gaboriaud C., Michon T., Kerfelec B., Chapus C.,					
RA Fontecilla-Camps J.C.;					
RT "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a					
RL highly structured truncated zymogen E.";					
EMBO J. 13:1763-1771(1994).					
-1 FUNCTION: DEFECTIVE ELASTASE-LIKE SERINE PROTEASE. DOES NOT SEEM					
TO HAVE A PROTEASE ACTIVITY. ITS LIKELY FUNCTION IS TO PROTECT					
PROCARBOXYPEPTIDASE A AGAINST DENATURATION IN THE ACIDIC					
ENVIRONMENT OF THE RUMINANT DUODENUM.					
-1 SUBUNIT: HETEROTRIMER OF SUBUNIT III; CARBOXYPEPTIDASE A AND					
CHYMOTRYPSINOGEN C.					
-1 SUBCELLULAR LOCATION: Extracellular.					
-1 TISSUE SPECIFICITY: PANCREAS.					
-1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.					
PIR; A25065; CPB0A3.					
PDB; 1FON; 14-OCT-96.					
MEROPS; S01.983; -.					
InterPro; IPR001314; Chymotrypsin.					
InterPro; IPR001254; Ser.protease_Try.					
Pfam; PF00089; trypsin; 1.					

```

DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SMO0020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SSR; 1.
DR KW PROPEP; propepase homolog; Pancreas; Digestion; 3D-structure.
DR FT PROPEP; 12 253 1
  FT CHAIN; 12 253 1
  FT DISULFID; 41 57
  FT DISULFID; 100 103
  FT DISULFID; 140 206
  FT DISULFID; 171 187
  FT DISULFID; 196 227
SQ SEQUENCE 253 AA; 27337 MW; 24663724D9A5409C CRC64;

Query Match      87.0%; Score 87; DB 1; Length 253;
Best Local Similarity 73.7%; Pred. No. 9.9e-07;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVNGSEAVPHKXWQVLSQ 19
DB 12 VVNGEDAVPSMSQVLSQ 30

RESULT 4
CTRA_BOVIN
ID_CTRA_BOVIN STANDARD; PRT; 245 AA.
AC P00766;
AT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Bovidae; Artiodactyla; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;
OC NCBI_TaxID:9913;
OX NCBI_TaxID:9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
  The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "The amino acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Kluh I., Kostka V., Moravsek L., Prusik Z., Vanacek J.,
  Keil B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie I.B., Hartley B.S.;
RT "Histidine sequences in the active centres of some 'serine'
  proteinases.";
RL Biochem. J. 101:232-241(1966).
RN [6]
RX X-RAY CRYSTALLOGRAPHY.
MEDLINE=72035052; PubMed=4399050;

```

Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
 "Serine Proteinases. The structure of alpha-chymotrypsin";
 Philo. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
 [7]

X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
 MEDLINE=70177557; PubMed=5442169.
 Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
 "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
 chymotrypsin, and implications for zymogen activation";
 Biochemistry 9:1997-2009(1970).
 [8]

X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
 MEDLINE=82078042; PubMed=6914398;
 Cohen G.H., Silverton E.W., Davies D.R.;
 "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
 Comparison with other pancreatic serine proteases";
 J. Mol. Biol. 148:443-475(1981).
 [9]

X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
 MEDLINE=86011575; PubMed=4046030,
 Tsukada H., Blow D.M.;
 "Structure of alpha-chymotrypsin refined at 1.68-A resolution";
 J. Mol. Biol. 184:703-711(1985).
 J. Mol. Biol. 184:703-711(1985).
 -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
 Phe-|-Xaa, Leu-|-Xaa.
 -|- SUBCELLULAR LOCATION: Extracellular.
 -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 -|- DATABASE: NAME=Worthington enzyme manual;
 WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
 PIR; AC0952; KY80A.
 DR PDB; 2CNA; 15-APR-90.
 DR PDB; 2CHN; 23-MAY-84.
 DR PDB; 4CHN; 23-OCT-84.
 DR PDB; 5CHA; 16-OCT-87.
 DR PDB; 6CHA; 16-OCT-87.
 DR PDB; 1CHG; 27-JAN-84.
 DR PDB; 1CHO; 16-JUL-88.
 DR PDB; 2GCH; 31-MAY-84.
 DR PDB; 3GCH; 15-OCT-92.
 DR PDB; 4GCH; 15-OCT-90.
 DR PDB; 5GCH; 15-OCT-90.
 DR PDB; 6GCH; 15-OCT-90.
 DR PDB; 7GCH; 15-OCT-90.
 DR PDB; 8GCH; 15-JUL-93.
 DR PDB; 1GCT; 15-OCT-91.
 DR PDB; 3GCT; 15-OCT-91.
 DR PDB; 1ACB; 31-OCT-91.
 DR PDB; 1GMC; 31-OCT-93.
 DR PDB; 1GMD; 31-OCT-93.
 DR PDB; 1CGI; 30-APR-94.
 DR PDB; 1CGJ; 30-APR-94.
 DR PDB; 1GCD; 22-JUN-94.
 DR PDB; 1GHA; 22-JUN-94.
 DR PDB; 1GHB; 22-JUN-94.
 DR PDB; 1GMH; 30-SEP-94.
 DR PDB; 2GMT; 01-NOV-94.
 DR PDB; 1MTN; 17-AUG-96.
 DR PDB; 1AB9; 20-AUG-97.
 DR PDB; 1FAO; 17-SEP-97.
 DR PDB; 1CAO; 23-JUL-97.
 DR PDB; 1YEA; 43-JUL-97.
 DR PDB; 1VGC; 12-NOV-97.
 DR PDB; 2VGC; 12-NOV-97.
 DR PDB; 3VGC; 12-NOV-97.
 DR PDB; 4VSC; 12-NOV-97.
 DR PDB; 1HJA; 14-JAN-98.
 MEROPS; S01.001; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR SMART; SMO0020; TRYP_SPC; 1.

DR PROSITE; P850240; TRYPSIN_DOM; 1.
DR PROSITE; P850134; TRYPSIN_HIS; 1.
DR PROSITE; P850135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW 3D-structure.
FT CHAIN 1 13 CHYMOTRYPSIN A, A CHAIN.
FT CHAIN 16 142 CHYMOTRYPSIN A, B CHAIN.
FT CHAIN 159 245 CHYMOTRYPSIN A, C CHAIN.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 1 122 CHARGE RELAY SYSTEM.
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 168 182
FT DISULFID 191 220
FT HELIX 12 15
FT STRAND 20 21
FT TURN 24 25
FT TURN 28 29
FT STRAND 30 34
FT STRAND 40 46
FT STRAND 51 54
FT TURN 57 58
FT TURN 62 63
FT STRAND 65 68
FT TURN 69 69
FT STRAND 72 72
FT TURN 73 74
FT STRAND 81 90
FT TURN 92 93
FT STRAND 95 95
FT TURN 96 99
FT STRAND 100 100
FT TURN 101 101
FT STRAND 104 108
FT STRAND 122 122
FT TURN 126 127
FT STRAND 132 133
FT STRAND 135 140
FT TURN 146 148
FT STRAND 154 154
FT STRAND 158 168
FT HELIX 169 172
FT HELIX 172 175
FT STRAND 180 184
FT TURN 192 193
FT TURN 195 196
FT STRAND 198 203
FT TURN 204 205
FT STRAND 206 216
FT TURN 218 219
FT TURN 222 223
FT STRAND 225 230
FT HELIX 231 233
FT STRAND 234 234
FT TURN 235 243
FT HELIX 235 243
SQ SEQUENCE 245 AA; 25666 MW; 91A9F28E2F3E3142 CRC64;

Query Match 85.0%; Score 85; DB 1; Length 245;
Best Local Similarity 84.2%; Pred. No. 26-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHKXKQNSLQ 19
| | | | | | | | | | | | | | | | | | | | |
Db 16 IVNGEAVPGSPHQNLSLQ 34

RESULT 5
EL3B HUMAN
ID EU3B HUMAN STANDARD; PRT; 270 AA.
AC P08861; P11423;

DT 01-NOV-1988 (Rel. 09, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase IIIB precursor (EC 3.4.21.70) (Protease E).
GN ELA3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=8807253; Pubmed=2826474;
RA Tani T., Ohsumi J., Mita K., Takiguchi Y.;
RT "Identification of a novel class of elastase isozyme, human
RT pancreatic elastase III, by cDNA and genomic gene cloning.";
RL J. Biol. Chem. 263:1231-1239(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-270 FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=8900943; Pubmed=3477287;
RA Shen W., Fletcher T.S., Lerman C.;
RT Primary structure of human pancreatic protease E determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:3447-3452(1987).
RN [4]
RP SEQUENCE OF 31-50.
RC TISSUE=Pancreas;
RX MEDLINE=89325560; Pubmed=2753124;
RA Moulard M., Kerfelec B., Mallet B., Chapus C.;
RT "Identification of a procarboxypeptidase A-truncated protease E
RT binary complex in human pancreatic juice.";
RL FEBS Lett. 250:166-170(1989).
RN [5]
RP SEQUENCE OF 94-164, AND CARBOHYDRATE-LINKAGE SITES ASN-114.
RC TISSUE=Pancreas;
RX MEDLINE=8928996; Pubmed=2737288;
RA Wendorf P., Geyer R., Szegoleit A., Linder D.;
RT "Localization and characterization of the glycosylation site of human
RT pancreatic elastase I.";
RL FEBS Lett. 249:275-278(1989).
CC -1- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
CC LITTLE ELASTOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Ala-|-Xaa. Does not
CC hydrolyse elastin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -1- CAUTION: Was originally (Ref.5) thought to be elastase I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M16630; AA036482.1;
DR EMBL; M16822; AA058454.1;
DR EMBL; M29934; B29934.
DR EPI; E27204; A27206.
DR EPI; S04999; S04999.
DR EPI; S04490; S04490.
DR HSP; P05805; IFO1.
DR MEROPS; S01.205; -.
DR GLYCOsuitedB; P08861; -.
DR SWISS-2DPAGE; P08861; HUMAN.
DR Genew; HGNC:15945; ELA3B.
DR

DR InterPro: IPR000114; Chymotrypsin.
 DR InterPro: IPR00135; Serine_protease_Try.
 DR Pfam: PF00689; trypsin. CHYMOTRYPSIN.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00202; TRYD_SPC.1
 DR PROSITE: PS50240; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER.1.
 KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
 FT SIGNAL 1 15 OR 16 (POTENTIAL).
 FT PROPEP 16 28 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 29 270 ELASTASE IIIB.
 FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 156 141 BY SIMILARITY.
 FT DISULFID 157 120 PROBABLE.
 FT DISULFID 157 220 BY SIMILARITY.
 FT DISULFID 188 264 BY SIMILARITY.
 FT DISULFID 213 244 N-LINKED GLYCOC. .).
 FT CARBOHYD 114 114 /FTID-CAR 000212. .).
 FT CONFLICT 4 4 R -> G (IN REF. 3).
 FT CONFLICT 64 64 A -> G (IN REF. 3).
 FT CONFLICT 79 79 W -> R (IN REF. 1).
 FT CONFLICT 129 131 MISSING (IN REF. 5).
 FT CONFLICT 164 164 R -> P (IN REF. 3).
 SQ SEQUENCE 270 AA; 22923 MW; B14BEOAAD3695AFE CRC64;
 Query Match 85.0%; Score 85; DB 1; Length 270;
 Best Local Similarity 73.1%; Pred. No. 2.2e-06;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 IYNGEEAVPHXKQWVLSQ 19
 DB 29 YNGEDAVFGSPWQVLSQ 47
 RESULT 6
 CTRB_BOVIN ID_CTRB_BOVIN STANDARD; PRT; 245 AA.
 AC P00767;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsinogen B (EC 3.4.21.1).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN 115_TaxID=9913;
 RP SEQUENCE DISULFIDE BONDS, AND ACTIVE SITE.
 RA MEDLINE=68218998; PubMed=5649671;
 RX Smillie L.B, Parkes A., Ragabathan N., Stevenson K.J., Parkes C.O.;
 RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and
 RT trypsinogen.";
 RL Nature 218:343-346(1968).
 CC NATURE 218:343-346(1968).
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
 CC Phe-|-Xaa, Leu-|-Xaa.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Consortium
 CC the European Bioinformatics Institute. There are no restriction on the
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL, K01173; AAA30841.1; .
 CC PIR: A21195; A21195.
 CC HSP: P00766; IACB.
 CC MEROPS: S01.152; .
 CC InterPro: IPR00134; Chymotrypsin.
 CC Pfam: PF00689; trypsin.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00202; TRYD_SPC.1
 CC InterPro: IPR00135; Serine_protease_Try.
 CC PROSITE: PS50240; TRYPSIN_DOM.1.
 CC PROSITE: PS00134; TRYPSIN_HIS.1.
 CC PROSITE: PS00135; TRYPSIN_SER.1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 FT SIGNAL 1 18 CHYMOTRYPSINOGE.2.
 FT CHAIN 19 263 CHYMOTRYPSINOGE.2.

SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 FT SIGNAL 1 18 CHYNOTRYPsinogen B.
 FT CHAIN 19 263 CHYNOTRYPsinogen B.
 FT CHAIN 19 31 CHYNOTRYPsin B, A CHAIN.
 FT CHAIN 34 164 CHYNOTRYPsin B, B CHAIN.
 FT CHAIN 167 263 CHYNOTRYPsin B, C CHAIN.
 FT ACT_SITE 75 75 CHARGE RELAY SYSTEM.
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 19 140 BY SIMILARITY.
 FT DISULFID 60 76 BY SIMILARITY.
 FT DISULFID 154 219 BY SIMILARITY.
 FT DISULFID 186 200 BY SIMILARITY.
 FT DISULFID 209 238 BY SIMILARITY.
 SQ SEQUENCE 263 AA; 27870 MW; 4C1C059A90B8701 CRC64;

Query Match 82.0%; Score 82; DB 1; Length 263;
 Best Local Similarity 78.9%; Pred. No. 6.4e-06;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXKQVSLQ 19
 DB 34 IVNGEDAVPGSPQVSLQ 52
 |||||:|||||
 |||||:|||||

RESULT 8
 ID CTRB_HUMAN STANDARD; PRT; 263 AA.
 AC P17538;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsinogen B precursor (SC 3.4.21.1).
 GN CTRB1 OR CTRB;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=99134264; PubMed=2917002;
 RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
 RA Mori T., Matsubara K.;
 RT "Molecular cloning and nucleotide sequence of human pancreatic
 RT prechymotrypsinogen cDNA."
 RL Biochem. Biophys. Res. Commun. 158:569-575 (1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
 CC Phe-|-Xaa, Leu-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).

CC EMBL; M24400; AA052128.1; -;
 CC EMBL; BC000385; AA05385.1; -;
 CC PIR; A31299; A31299.
 CC HSSP; P00766; 1CHG.
 CC MEROPS; S01.152; -;
 CC Genew; HGNC:2521; CTRB1.
 CC MIM; 118890; IP001114; Chymotrypsin.
 CC InterPro; IP001254; Ser_Protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYNOTRYPsin.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00134; TRYPSIN_DOM; 1.
 CC PROSITE; PS00135; TRYPSIN_HIS; 1.

modified and this statement is not removed. Usage by and for commercial

RESULT 12

CLCR HUMAN STANDARD; PRT; 268 AA.

ID CLCR HUMAN STANDARD; PRT; 268 AA.

AC Q99895; Q9NHS; Q00765;

AD 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Caldecrin precursor (EC 3.4.21.2) (Chymotrypsin C).

GN CTRC OR CLCR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxId=9606;

[1]

SEQUENCE FROM N.A., AND VARIANT TRP-80.

RC TISSUE=Pancreas;

RC MEDLINE=96211265; PubMed=8635596;

RX Tomomura A., Akiyama M., Itoh H., Yoshino I., Tomomura M., Nishii Y.,

RA Nolkura T., Sanehi T.; expression of human caldecrin.;

RT "Molecular cloning and expression of human caldecrin.;"

RL PNAS Lett. 386:26-28(1996).

[2]

SEQUENCE FROM N.A.

RP Coville G.;

RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RL [3]

SEQUENCE OF 17-268 FROM N.A.

RC TISSUE=Pancreas;

RC MEDLINE=96211265; PubMed=8635596;

RX Tomomura A., Akiyama M., Itoh H., Yoshino I., Tomomura M., Nishii Y.,

RA Nolkura T., Sanehi T.; expression of human caldecrin.;

RT "Molecular cloning and expression of human caldecrin.;"

RL PNAS Lett. 386:26-28(1996).

[4]

CHARACTERIZATION.

RP MEDLINE=98207038; PubMed=9538241;

RX Yoshino-Yasuda I., Kobayashi K., Akiyama M., Itoh H., Tomomura A.,

RA Sanehi T.;

RT "Caldecrin is a novel-type serine protease expressed in pancreas, but

RT its homologue, elastase IV, is an artifact during cloning derived

RT from caldecrin gene.;"

RL J. Biochem. 123:546-554(1998).

[5]

FUNCTION: HAS CHYMOTRYPSIN-TYPE PROTEASE ACTIVITY AND HYPOCALCEMIC

ACTIVITY.

-1- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa,

Phe-|-Xaa, Met-|-Xaa, Trp-|-Xaa, Glu-|-Xaa, Asn-|-Xaa.

-1- TISSUE SPECIFICITY: PANCREAS.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; S62198; AAB47104.2; ALT_SEQ.

EMBL; AL031283; CAB77355.1; "-

EMBL; Y13697; CAA74031.1; "-

HSSP; P00766; 1CHG.

MEROPS; S01.157; "-

Genew; HGNC:2523; CTRC.

MIM; 601405; "-

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR001254; Ser_protease_Try.

Pfam; PF00089; trypsin; 1.

PRINTS; PRO0722; CHYMOTRYPSIN.

SMART; SM00020; TRYD_SPC; 1.

PROSITE; PS00240; TRYPSIN_DOM; 1.

PROSITE; PS00134; TRYPSIN_HTS; 1.

PROSITE; PS00135; TRYPSIN_SER; 1.

Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW Polymorphism. 1 15
FT SIGNAL 17 29
FT PROPEP 1 29
FT CHAIN 30 268
FT ACT_SITE 74 74
FT ACT_SITE 121 121
FT ACT_SITE 216 216
FT DISULFID 17 141
FT DISULFID 17 141
FT DISULFID 159 222
FT DISULFID 159 222
FT DISULFID 183 242
FT DISULFID 212 242
FT CARBOHYD 25 52
FT CARBOHYD 25 52
FT CARBOHYD 226 226
FT VARIANT 80 80
FT CONFLICT 16 16
FT CONFLICT 52 52
SQ SEQUENCE 268 AA; 29484 MM; 46DBF3B4A96516F CRC64;
Query Match 77.0%; Score 77; DB 1; Length 268;
Best Local Similarity 63.2%; Pred. No. 4e-05;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 IVNGEAVPHWKNQVSLQ 19
DB 30 VVGGEARNPSWQVSLQ 48
:|||||:|||||
RESULT 13
ELZA HUMAN STANDARD; PRT: 269 AA.
ID ELZA_HUMAN
AC P08217; Q14243;
DC 01-AUG-1988 (Rel. 08, Created)
DT 15-JUN-2005 (Rel. 05, Last annotation update)
DE Elastase 2A precursor (EC 3.4.21.71).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RN cDNAs are expressed in human pancreas.";
RL [2] 6:165-172 (1987).
RP SEQUENCE FROM N.A.
RX MEDLINE=8107669; PubMed=3427074;
RA Fletcher T.S., Shen W.F., Lorgman C.;
RT "Primary structure of human pancreatic elastase 2 determined by
RN sequence analysis of the cloned mRNA.";
RL [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89198076; PubMed=28934346;
RA Shirasu Y., Yoshida H., Matsuki S.,
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RN encoding human pancreatic elastase 2.";
RL [4]
RP SEQUENCE FROM N.A.
RX Thomae D.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL [5]
RP SEQUENCE FROM N.A.

Search completed: February 12, 2003, 10:23:15
Job time : 3.68657 secs


```

Query Match      95.0%; Score 95; DB 13; Length 263;
Best Local Similarity 89.5%; Pred No. 6.3e-08;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 1VNGEEAVPHXKXQVSLQ 19
DB 32 1VNGEEAVPHSPWQVSLQ 50

RESULT 2
ID Q9D7T9 PRELIMINARY; PRT; 269 AA.
AC Q9D7T9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2310074FOIRik protein.
DE ELA3B OR 2310074FOIRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schrim L.M., Scubili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okado T., Furuno M., Anno H., Baldarelli R., Barab G.,
RA Blake J., Raffalli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guncich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momotani N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
RA Hayashizaki Y.
RT Functional annotation of a full-length mouse cDNA collection.*;
RL Nucleotide 409,685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK008858; BAB25932.1; -.
DR HSSP; P05805; 1FON.
DR MGd; MG1;1915118; Elab3b.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KX Hydroxylase; Serine protease.
SQ SEQUENCE 269 AA; 28977 MW; 9F43F769DDB2A7CF CRC64;

Query Match      94.0%; Score 94; DB 11; Length 269;
Best Local Similarity 84.2%; Pred. No. 6.3e-08;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 1VNGEEAVPHXKXQVSLQ 19
DB 28 1VNGEEAVPHSPWQVSLQ 46

RESULT 4
ID Q9WTQ3 PRELIMINARY; PRT; 260 AA.
AC Q9WTQ3
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Chymotrypsinogen 2.
OS Paralichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acantopterygii; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuroncti; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC NCBI_TaxID=98255;

[1]_TaxID=98255;

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA Suzuki T., Silvestava A.S., Kurokawa T.;

RL "Japanese flounder mRNA for chymotrypsinogen 2.";

RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC -!- TRYPSIN FAMILY.

DR EMBL; A029754; BA82366.1; -.

DR M567; A00766; ICMG.

DR MEROPS; S01152;

DR InterPro; IPR001354; Chymotrypsin.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00689; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; tryp_SPC; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 92.0%; Score 92; DB 13; Length 260;

Best Local Similarity 84.2%; Pred. No. 1.3e-07;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVNGEAVPHXKXWQVSLQ 19

DB 31 LVNGEAVPHXKXWQVSLQ 49

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 6

O19023

ID O19023

AC O19023;

PRELIMINARY;

PRT; 257 AA.

DT 01-JAN-1998 (TRENBLrel. 05, Created)

DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

GN Elastase (EC 3.4.21.36) (Fragment).

ELAI.

OC Macaca mulatta (Rhesus macaque).

OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cebopithecoidea; Macaca.

OC NCBI_TaxID=9544;

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA "Pangreatic elastase from rhesus monkey.";

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC -!- TRYPSIN FAMILY.

DR EMBL; AJ000067; CA03899.1; -.

DR HSP; P05805; IFON.

DR MEROPS; S01154; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00689; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; tryp_SPC; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

FT NON_TER 1

FT NON_TER 257

SQ SEQUENCE 257 AA; 27687 MW; 4D443DB67233D8DC CRC64;

Query Match 85.0%; Score 85; DB 6; Length 257;

Best Local Similarity 73.7%; Pred. No. 2e-06;

Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVNGEAVPHXKXWQVSLQ 19

DB 16 VNGEAVPHXKXWQVSLQ 34

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acantopterygii; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuroncti; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC NCBI_TaxID=98255;

[1]_TaxID=98255;

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA Suzuki T., Silvestava A.S., Kurokawa T.;

RL "Japanese flounder mRNA for chymotrypsinogen 2.";

RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC -!- TRYPSIN FAMILY.

DR EMBL; A029754; BA82366.1; -.

DR M567; A00766; ICMG.

DR MEROPS; S01152;

DR InterPro; IPR001354; Chymotrypsin.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00689; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; tryp_SPC; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 92.0%; Score 92; DB 13; Length 260;

Best Local Similarity 84.2%; Pred. No. 1.3e-07;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVNGEAVPHXKXWQVSLQ 19

DB 31 LVNGEAVPHXKXWQVSLQ 49

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acantopterygii; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuroncti; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC NCBI_TaxID=98255;

[1]_TaxID=98255;

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA Suzuki T., Silvestava A.S., Kurokawa T.;

RL "Japanese flounder mRNA for chymotrypsinogen 2.";

RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC -!- TRYPSIN FAMILY.

DR EMBL; A029754; BA82366.1; -.

DR M567; A00766; ICMG.

DR MEROPS; S01152;

DR InterPro; IPR001354; Chymotrypsin.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00689; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; tryp_SPC; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 92.0%; Score 92; DB 13; Length 260;

Best Local Similarity 84.2%; Pred. No. 1.3e-07;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVNGEAVPHXKXWQVSLQ 19

DB 31 LVNGEAVPHXKXWQVSLQ 49

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

```

RL Nature 409:685-690 (2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL AK007566, BAB25112.1; -.
DR HSP: P00766; IGCT.
DR MEROPS: S01.152; -.
DR MGD; MG1:1913723; Z20008D09Rik.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; SM00722; CHYMOTRYPSIN.
DR SMART; SM0020; TRYPSIN_1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease.
KW SEQUENCE 283 AA; 27898 MW; C0638FB8F905A92F CRC64;
SQ Query Match 81.0%; Score 81; DB 11; Length 263;
Best Local Similarity 73.7%; Pred No. 9, 9e-06;
Indels 0; Gaps 0;

```

Q9CR35
TO C0C015
PRELIMINARY: PRT: 263 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 2200008D09rik protein.

SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
 RX MEDLINE=21085660; PubMed=11217851; Yoshino M., Itoh M., Iehi Y.,
 RA Kawai J., Shinagawa N., Shibata K., Konno H., Adachi J., Fukuda S.,
 RA Arakawa K., Hara A., Nishikuni Y., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saizawa T., Izawa M., Kikuchi T., Bono H., Kasukawa T., Saito R.,
 RA Saio T., Okazaki H., Gojobori T., Apburner M., Batalov S., Casavant T.,
 RA Kadota K., Masuda H.A., Apburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuda Y., Nikaido I., Pease G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Okido T., Furuno M., Anono H., Baldarelli R., Barah G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Masima Y., Mazzarelli J., Momabato N.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weiss C., Whitaker C., Wilmig L.,
 RA Wynehan-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtauki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001);
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY.
 CC EMBL; AK008927; BAB25971.1; -
 DR EMBL; AK003079; BAB2553.1; -
 DR EMBL; AK007765; BAB2541.1; -
 DR EMBL; AK007815; BAB25480.1; -
 DR EMBL; AK008729; BAB25861.1; -
 DR EMBL; AK008868; BAB25954.1; -
 DR HSSP; P00766; IGCT

```

DR MEROPS; S01.152;
DR MGD; MG1:1913723; 2200080D09Rik.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR SMART; SMO00722; CHYMOTRYPSIN.
DR SMART; SMO0020; trypsin; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR HYDROLASE; Serine protease.
DR SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match
Best Local Similarity 81.0%; Score 81; DB 11; Length 263;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
DB 34 IVNGEAVPGSNPMQVSLQ 52

RESULT 10
QY0960 PRELIMINARY; PRT; 264 AA.
AC QY0960;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE 1810004D158IK protein.
GN CTRL OR 1810004D158IK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Peole G., Quackenbush J.,
RA Schrim L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Ring B., Ringwald M., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Nordone P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nucleic Acids Res. 30(12):2001.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007333; BAB24967.1; -.
DR HSSP; P00766; 4CHA.
DR MGD; MG1:89558; Ctrl.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF000722; Ser_protease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR HYDROLASE; Serine protease.
DR SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match
Best Local Similarity 80.0%; Score 80; DB 11; Length 264;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
DB 34 IVNGEAVPGSNPMQVSLQ 52

```

```

RESULT 12
O95Q28 PRELIMINARY; PRT; 264 AA.
AC Q95Q28;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin, legicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Soqane Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yanaguchi N.;
RL "Molecular cloning of rat chymopasin.";
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
EMBL; AB020757; BAB20287.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
AC Q95Q28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to elastase 3, pancreatic (protease E).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
EMBL; BC008383; AAH08383.1; -.
DR MEROPS; S01.154; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 264 AA; 2816 MW; 395D5D210F03500E CRC64;

Query Match 80.0%; Score 80; DB 11; Length 264;
Best Local Similarity 78.9%; Pred No. 1.5e-05;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IYNGEENPHXKXWQVSLQ 19
Db 34 IYNGENAVFGSPWQVSLQ 52

RESULT 14
O95Q18 PRELIMINARY; PRT; 270 AA.
AC Q95Q18;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to elastase 3, pancreatic (protease E).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
EMBL; BC008383; AAH08383.1; -.
DR MEROPS; S01.154; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 270 AA; 2946 MW; 946DDBA694A102E CRC64;

Query Match 80.0%; Score 80; DB 4; Length 270;
Best Local Similarity 66.4%; Pred No. 1.5e-05;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IYNGEENPHXKXWQVSLQ 19
Db 29 VVHGDNVPSWQVSLQ 47

RESULT 15
O8TAN2 PRELIMINARY; PRT; 461 AA.
AC O8TAN2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Midgut serine proteinase-3.
OS Rhinipcephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhinipcephalus.
OX NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Mulenga A., Onuma M., Sugimoto C.;
RT "Rhipicephalus appendiculatus midgut serine proteinase-3, cDNA cloning
and characterization.";
PI Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL:AF078095; AAL79567.1; -. 641CDCA5CF2EB68 CRC64;
DB EMBL:AF078095; AAL79567.1; -. 641CDCA5CF2EB68 CRC64;
SQ SEQUENCE 461 AA; 43561 MW; 641CDCA5CF2EB68 CRC64;

Query Match 77.0%; Score 77; DB 5; Length 461;
Best Local Similarity 68.4%; Pos. No. 8.6e-05;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 IVNGEEAVPHXXKQVSLQ 19
|||:|||||:|:
Db 208 IVAGQEAIPHSMQASVQ 226

Search completed: February 12, 2003, 10:27:23
Job time : 15.1493 secs

THIS PAGE BLANK (uspro)


```

/ CLASSIFICATION:
/ PRIOR APPLICATION DATA: US 08/363,092
/ APPLICATION NUMBER: 20-DEC-1994
/ FILING DATE: 20-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Savitzky, Martin
/ REGISTRATION NUMBER: 29,699
/ REFERENCE/DOCKET NUMBER: A1172-WO
/ TELEPHONE: (610)454-3816
/ TELEFAX: (610)454-3808
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 31 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Sus scrofa
PCT-US95-16826-1

Query Match 85.0%; Score 85; DB 5; Length 31;
Best Local Similarity 73.7%; Pred. No. 2.7e-08;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVNGEAVPHYKXWQVSLQ 19
:|||||:|:|:|
Db 1 VVNGEDAVPYSPWQVSLQ 19

RESULT 2
PCT-US95-16826-2
/ Sequence 2, Application PC/TUS9516826
/ GENERAL INFORMATION:
/ APPLICANT: NAME: OSTEOSA, INC.
/ APPLICANT: STREET: 2040 Babcock Road, Suite 201
/ APPLICANT: CITY: San Antonio
/ APPLICANT: STATE: Texas
/ APPLICANT: COUNTRY: USA
/ APPLICANT: POSTAL CODE: 78229
/ APPLICANT:
/ APPLICANT: NAME: YONEDA, Toshiyuki
/ APPLICANT: STREET: 3530 Hunter's Sound
/ APPLICANT: CITY: San Antonio
/ APPLICANT: STATE: Texas
/ APPLICANT: COUNTRY: USA
/ APPLICANT: POSTAL CODE: 78230
/ APPLICANT:
/ APPLICANT: NAME: IZBIKA, Elzbieta
/ APPLICANT: STREET: 7738 Apple Green
/ APPLICANT: CITY: San Antonio
/ APPLICANT: STATE: Texas
/ APPLICANT: COUNTRY: USA
/ APPLICANT: POSTAL CODE: 78240
/ APPLICANT:
/ APPLICANT: NAME: MUNDY, Gregory R.
/ APPLICANT: STREET: 3719 Morgan's Creek
/ APPLICANT: CITY: San Antonio
/ APPLICANT: STATE: Texas
/ APPLICANT: COUNTRY: USA
/ APPLICANT: POSTAL CODE: 78230
/ APPLICANT:
/ TITLE OF INVENTION: Modulators of Bone Cell Function and
/ NUMBER OF SEQUENCES: 2
/ TITLE OF INVENTION: Uses Thereof
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Rd. 3C43
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19002

. . .

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/16826
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/363,092
/ FILING DATE: 20-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Savitzky, Martin
/ REGISTRATION NUMBER: 29,699
/ REFERENCE/DOCKET NUMBER: A1172-WO
/ TELEPHONE: (610)454-3816
/ TELEFAX: (610)454-3808
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
PCT-US95-16826-2

Query Match 85.0%; Score 85; DB 5; Length 32;
Best Local Similarity 73.7%; Pred. No. 2.8e-08;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IVNGEAVPHYKXWQVSLQ 19
:|||||:|:|:|
Db 1 VVNGEDAVPYSPWQVSLQ 19

RESULT 3
US-08-278-091-10
/ Sequence 10, Application US/08278091
/ Patent No. 5506139
/ GENERAL INFORMATION:
/ APPLICANT: LOOMORE, Sheena M
/ APPLICANT: YANG, Yan-Ping
/ APPLICANT: CHONG, Pele
/ APPLICANT: COMEN, Raymond P.
/ APPLICANT: KLEIN, Michel H.
/ TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
/ NUMBER OF SEQUENCES: 23
/ TITLE OF INVENTION: Reduced Protease Activity
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: Suite 701, 330 University Avenue
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US
/ FILING DATE: 21-JUL-1994
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Savitzky, Martin
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-371
/ TELECOMMUNICATION INFORMATION:
```


TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-278-091-10
 Query Match 85.0%; Score 85; DB 1; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
 Db 1 IVNGEAVPGSMPQVSLQ 19

RESULT 4
 US-08-483-859-10

Sequence 10, Application US/08483859
 Patent No. 5656436
 GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: COHEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 NUCLEOTIDE SEQUENCES: 23
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Sam & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/483,859
 FILING DATE: 26-AUG-1994
 CLASSIFICATION: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/296,149
 FILING DATE: 26-AUG-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/278,091
 FILING DATE: 21-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match 85.0%; Score 85; DB 1; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
 Db 1 IVNGEAVPGSMPQVSLQ 19

RESULT 5

US-08-472-173-10
 Sequence 10, Application US/08472173
 Patent No. 5665353
 GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: COHEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
 NUCLEOTIDE SEQUENCES: 23
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Sam & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,173
 FILING DATE: 26-AUG-1994
 CLASSIFICATION: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/296,149
 FILING DATE: 26-AUG-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/278,091
 FILING DATE: 21-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match 85.0%; Score 85; DB 1; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
 Db 1 IVNGEAVPGSMPQVSLQ 19

RESULT 6

US-08-487-167-10
 Sequence 10, Application US/08487167
 Patent No. 5669302
 GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: COHEN, Raymond P.

us-10-036-371-4-rai

Wed Feb 12 11:59:37 2003

```

APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
REDUCED PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,167
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-482-816-10

Query Match 85.0%; Score 85; DB 2; Length 228;
Best Local Similarity 84.2%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEAVPHXKXWQVSLQ 19
| | | | | | | | | | | | | | | | | | | |
DB 1 IVNGEAVPGSPWQVSLQ 19

RESULT 7
US-08-296-149-10
Sequence 10, Application US/08296149
Patent No. 5939297
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: COMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
REDUCED PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,149
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-390
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-296-149-10

Query Match      85.0%; Score 85; DB 2; Length 228;
Best Local Similarity 84.2%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKWQVSLQ 19
Db 1 IVNGEEAVPGSWPQVSLQ 19

RESULT 9
US-08-801-499-10
; Sequence 10, Application US/08801499
; Patent No. 5962430
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; COUNTRY: Ontario
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,499
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/482,816
; FILING DATE: 07-JUN-1995
; APPLICATION DATA:
; FILING DATE: 08/296,149
; PRIOR APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 1038-671 MTS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-801-499-10

Query Match      85.0%; Score 85; DB 2; Length 228;
Best Local Similarity 84.2%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKWQVSLQ 19
Db 1 IVNGEEAVPGSWPQVSLQ 19

RESULT 10
US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-615-271-10

Query Match      85.0%; Score 85; DB 2; Length 228;
Best Local Similarity 84.2%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMKWQVSLQ 19
Db 1 IVNGEEAVPGSWPQVSLQ 19

RESULT 11
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto

```

Wed Feb 12 11:59:37 2003

```

; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-074-659-10
;
; Query Match 85.0%; Score 85; DB 3; Length 228;
; Best Local Similarity 84.2%; Pred. No. 2.5e-07;
; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; Qy 1 IVNGEAVPHYKXWQVSLQ 19
; | | | | | | | | | | | | | | | | | |
; Db 1 IVNGEAVPGSNWQVSLQ 19
;
; RESULT 12
; US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,659
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-730 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-074-659-10
;
; Query Match 85.0%; Score 85; DB 3; Length 228;
; Best Local Similarity 84.2%; Pred. No. 2.5e-07;
; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; Qy 1 IVNGEAVPHYKXWQVSLQ 19
; | | | | | | | | | | | | | | | | | |
; Db 1 IVNGEAVPGSNWQVSLQ 19
;
; RESULT 13
; US-09-106-468-10
; Sequence 10, Application US/09106468
; Patent No. 6114125
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,468
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155

```

TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-106-468-10

Query Match 85.0%; Score 85; DB 3; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXKXWQVSLQ 19
 |||||
 Db 1 IVNGEEAVPGSNPWQVSLQ 19

RESULT 14
 US-09-106-466A-10
 Sequence 10, Application US/09106466A
 Patent No. 6147057

GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: OMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: ANALOG OF HARMOPHILUS HIN47 WITH REDUCED
 NUMBER OF SEQUENCES: 23
 TITLE OF INVENTION: PROTEASE ACTIVITY

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 6th Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada

ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/106,466A

FILING DATE:
 CLASSIFICATION: 514514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/615,271

FILING DATE: 20-JUN-1996
 CLASSIFICATION: 514514
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1039-826
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-106-466A-10

Query Match 85.0%; Score 85; DB 4; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXKXWQVSLQ 19
 |||||
 Db 1 IVNGEEAVPGSNPWQVSLQ 19

RESULT 15

US-09-106-467-10
 Sequence 10, Application US/09106467
 Patent No. 6153580

GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: OMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
 NUMBER OF SEQUENCES: 23
 TITLE OF INVENTION: PROTEASE ACTIVITY

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 6th Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada

ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/106,467
 FILING DATE:
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/615,271
 FILING DATE: 20-JUN-1996

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-824
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 228 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-106-467-10

Query Match 85.0%; Score 85; DB 4; Length 228;
 Best Local Similarity 84.2%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPHXKXWQVSLQ 19
 |||||
 Db 1 IVNGEEAVPGSNPWQVSLQ 19

Search completed: February 12, 2003, 10:30:09
 Job time : 5.29851 secs

THIS PAGE BLANK (USPTO)

1	95	95.0	20	12	US-10-036-371-4	Sequence 4, Appl.
2	85	85.0	20	12	US-10-036-371-6	Sequence 6, Appl.
3	82	82.0	146	10	US-09-925-297-643	Sequence 643, Appl.
4	82	82.0	263	10	US-09-888-615-96	Sequence 96, Appl.
5	80	80.0	270	10	US-09-923-779-152	Sequence 152, Appl.
6	78	78.0	273	10	US-09-925-297-695	Sequence 695, Appl.
7	74	74.0	269	10	US-09-925-297-576	Sequence 576, Appl.
8	70	70.0	192	10	US-09-925-297-629	Sequence 629, Appl.
9	70	70.0	970	10	US-09-888-615-101	Sequence 101, Appl.
10	69	69.0	113	10	US-09-925-297-871	Sequence 871, Appl.
11	66	66.0	283	10	US-09-988-976A-1	Sequence 1, Appl.
12	66	66.0	384	9	US-09-981-353-23	Sequence 23, Appl.
13	66	66.0	393	9	US-10-012-896-94	Sequence 94, Appl.
14	66	66.0	393	9	US-09-895-793-94	Sequence 934, Appl.
15	66	66.0	393	9	US-09-895-814-94	Sequence 934, Appl.
16	66	66.0	393	10	US-09-759-143-934	Sequence 934, Appl.
17	66	66.0	393	10	US-09-759-143-934	Sequence 934, Appl.
18	66	66.0	393	10	US-09-780-669-934	Sequence 934, Appl.
19	66	66.0	393	10	US-09-822-827-934	Sequence 934, Appl.
20	66	66.0	492	9	US-10-012-896-895	Sequence 895, Appl.

```

Query Match      95.0%; Score 95; DB 12; Length 20;
Eest Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPHXKXQVSLQ 19
Db 1 IVNGEAVPHXKXQVSLQ 19

```

```

; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-643

Query Match      82.0%; Score 82; DB 10; Length 146;
Best Local Similarity 78.9%; Pred. No. 2.5e-06;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHKXKQVLSQ 19
Db 38 IVNGEDAVPGSMQVLSQ 56

RESULT 4
US-09-888-615-96
; Sequence 96, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-96

Query Match      82.0%; Score 82; DB 10; Length 263;
Best Local Similarity 78.9%; Pred. No. 4.6e-06;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHKXKQVLSQ 19
Db 34 IVNGEDAVPGSMQVLSQ 52

RESULT 5
US-09-923-779-152
; Sequence 152, Application US/09923779
; Patent No. US2002007621A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923,779
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-779-152

Query Match      80.0%; Score 80; DB 10; Length 270;
Best Local Similarity 88.4%; Pred. No. 9.8e-06;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

US-10-036-371-6
; Sequence 643, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BUARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; COSMETIC USE
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/411,688
; PRIOR APPLICATION NUMBER: 086/99
; PRIOR FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH 20
; TYPE: PRT
; ORGANISM: Bovine sp.
US-10-036-371-6

Query Match      85.0%; Score 85; DB 12; Length 20;
Best Local Similarity 84.2%; Pred. No. 1.1e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHKXKQVLSQ 19
Db 1 IVNGEDAVPGSMQVLSQ 19

RESULT 3
US-09-925-297-643
; Sequence 643, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL05
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 643
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (121)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```



```

Qy 1 IVNGEEAVPHXKXQVSLQ 19
Db 29 VVHGDAVPYSPMQVSLQ 47

RESULT 6
US-09-925-297-695
; Sequence 695, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-695

Query Match 78.0%; Score 78; DB 10; Length 273;
Best Local Similarity 69.4%; Pred. No. 2,4e-05;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXKXQVSLQ 19
Db 32 VVHGDAVPYSPMQVSLQ 50

RESULT 7
US-09-925-297-576
; Sequence 576, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 576
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

```

; NAME/KEY: SITE
; LOCATION: (213)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (220)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-576

Query Match 74.0%; Score 74; DB 10; Length 269;
Best Local Similarity 68.4%; Pred. No. 9e-05;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXKXQVSLQ 19
Db 34 VVGEERAPNSPFWQVSLQ 52

RESULT 8
US-09-925-297-529
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 529
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-297-529

Query Match 70.0%; Score 70; DB 10; Length 192;
Best Local Similarity 73.7%; Pred. No. 0.00028;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXKXQVSLQ 19
Db 30 IVNGENAVLGSPPQVSLQ 48

RESULT 9
US-09-888-615-101
; Sequence 101, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYCZAK, GLEN
; APPLICANT: HANING, GERARD
; APPLICANT: SUDAN, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/121
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 970
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-09-888-613-101

Query Match      70.0%; Score 70; DB 10; Length 970;
Best Local Similarity 63.2%; Pred. No. 0.0015;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMXQVSLQ 19
Db 433 IAGGEACFCHPWQVGLR 451

RESULT 10
US-09-925-297-871 Application US/09925297
; Sequence 871; Patent No. US20020119531A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P1105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 871
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (70)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (87)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; SEQ ID NO 871
Query Match      69.0%; Score 69; DB 10; Length 113;
Best Local Similarity 63.2%; Pred. No. 0.00023;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMXQVSLQ 19
Db 29 MCGEEARPNPQVSLQ 47

RESULT 11
US-09-988-975A-1
; Sequence 1; Application US/09988975A
; Patent No. US20020119531A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: PROSTATE-ASSOCIATED PROTEASE ANTIBODY
; FILE REFERENCE: PF-0227-2 CIP
; CURRENT APPLICATION NUMBER: US/09/988,975A
; CURRENT FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PERL Program
; SEQ ID NO 1

```

```

; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020119531A1 556016
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 235
; OTHER INFORMATION: unknown or other
US-09-988-975A-1

Query Match      66.0%; Score 66; DB 10; Length 283;
Best Local Similarity 66.7%; Pred. No. 0.0018;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMXQVSL 18
Db 47 IVGGESALPGAMPQVSL 64

RESULT 12
US-09-981-353-23
; Sequence 23; Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: P1098 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 991163CD1
US-09-981-353-23

Query Match      66.0%; Score 66; DB 9; Length 384;
Best Local Similarity 66.7%; Pred. No. 0.0025;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPHXMXQVSL 18
Db 148 IVGGESALPGAMPQVSL 165

RESULT 13
US-10-012-896-934
; Sequence 934; Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Ruqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Patrick
; APPLICANT: Wang, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henpler, William T.
; APPLICANT: Henderson, Robert A.

```

```

; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: McNeill, John
; APPLICANT: Mantanabara, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-934

```

```

Query Match          66.0%; Score 66; DB 9; Length 393;
Best Local Similarity 66.7%; Pred. No. 0.0025; 5; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 5;

```

```

QY 1 IVNGEAVPHKXKQVSL 18
    |||||:|||||
Db 157 IVGESALPGAMPQVSL 174

```

```

RESULT 14
US-09-895-793-934
; Sequence 934, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-934

```

```

Query Match          66.0%; Score 66; DB 9; Length 393;
Best Local Similarity 66.7%; Pred. No. 0.0025; 5; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 5;

```

```

QY 1 IVNGEAVPHKXKQVSL 18
    |||||:|||||
Db 157 IVGESALPGAMPQVSL 174

RESULT 15
US-09-895-814-934
; Sequence 934, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-934

```

```

Query Match          66.0%; Score 66; DB 9; Length 393;
Best Local Similarity 66.7%; Pred. No. 0.0025; 5; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 5;

```

```

QY 1 IVNGEAVPHKXKQVSL 18
    |||||:|||||
Db 157 IVGESALPGAMPQVSL 174

```

```

Search completed: February 12, 2003, 10:31:05
Job time : 3.0447s secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM protein - protein search, using sw model
 Run on: February 12, 2003, 10:03:55 ; Search time 8.73134 Seconds
 without alignments
 198.395 Million cell updates/sec

Title: US-10-036-371-5
 Perfect score: 68
 Sequence: 1 CQVPAIQPVLSGL 13
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
 Total number of hits satisfying chosen parameters: 908470
 Minimum DB seq length: 0
 Maximum DB seq length: 200000000
 Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database : A_Geneseq_101002.*
 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	13	AA50209	Neutrophil-activat
2	68	100.0	13	AA131577	Amino terminal seq
3	68	100.0	22	AA50207	Neutrophil-activat
4	68	100.0	24	AA50207	Neutrophil-activat
5	63	92.6	146	AA199596	Bovine chymotryps
6	63	92.6	263	AA199596	Human pancreatic c
7	62	91.2	263	AA199596	Amino acid sequenc
8	62	91.2	15	AA50208	Neutrophil-activat
9	49	72.1	15	AA50210	Neutrophil-activat
10	49	72.1	192	AA198477	Human pancreatic c
			264	AA11710	Human serine prote

11	48	70.6	264	21	AA11711	Mouse serine prote
12	46	67.6	13	40	AA50212	Neutrophil-activat
13	42	61.8	13	40	AA50211	Neutrophil-activat
14	42	61.8	107	22	AA50211	Neutrophil-activat
15	42	61.8	107	22	AA50211	Neutrophil-activat
16	42	61.8	144	22	AA50211	Neutrophil-activat
17	42	61.8	170	22	AA50211	Neutrophil-activat
18	42	61.8	170	22	AA50211	Neutrophil-activat
19	42	61.8	327	20	AA50211	Neutrophil-activat
20	42	61.8	327	20	AA50211	Neutrophil-activat
21	42	61.8	339	22	AA50211	Neutrophil-activat
22	42	61.8	339	22	AA50211	Neutrophil-activat
23	42	61.8	339	22	AA50211	Neutrophil-activat
24	40	58.8	13	21	AA50211	Neutrophil-activat
25	40	58.8	13	21	AA50211	Neutrophil-activat
26	39	57.4	13	19	AA50211	Neutrophil-activat
27	39	57.4	13	19	AA50211	Neutrophil-activat
28	39	57.4	91	22	AA50211	Neutrophil-activat
29	39	57.4	168	22	AA50211	Neutrophil-activat
30	39	57.4	308	22	AA50211	Neutrophil-activat
31	39	57.4	319	22	AA50211	Neutrophil-activat
32	39	57.4	320	22	AA50211	Neutrophil-activat
33	39	57.4	320	22	AA50211	Neutrophil-activat
34	39	57.4	320	22	AA50211	Neutrophil-activat
35	39	57.4	323	22	AA50211	Neutrophil-activat
36	39	57.4	323	22	AA50211	Neutrophil-activat
37	39	57.4	327	22	AA50211	Neutrophil-activat
38	39	57.4	328	22	AA50211	Neutrophil-activat
39	39	57.4	330	22	AA50211	Neutrophil-activat
40	39	57.4	333	22	AA50211	Neutrophil-activat
41	39	57.4	334	22	AA50211	Neutrophil-activat
42	39	57.4	339	22	AA50211	Neutrophil-activat
43	39	57.4	339	22	AA50211	Neutrophil-activat
44	39	57.4	525	15	AA50211	Neutrophil-activat
45	39	57.4	609	22	AA50211	Neutrophil-activat

ALIGNMENTS

RESULT 1
 ID AA50209 standard; Peptide; 13 AA.
 XX
 XX AA50209;
 DT 12-JAN-2000 (first entry)
 XX
 DE Neutrophil-activating pancreatic derived peptide 9.
 KW Cell activation; pancreas; treatment; cardiovascular disease; trauma;
 KW inflammatory disease; autoimmune disease; arthritis; diabetes; stroke;
 KW organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
 KW hemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
 KW trauma; protease inhibitor; hypertension; sepsis.
 XX
 OS Bos taurus.
 XX
 XX WO9946367-A2.
 XX
 XX 16-SEP-1999.
 XX
 PF 11-MAR-1999; 99WO-US05247.
 XX
 PR 11-MAR-1999; 98US-0038894.
 XX
 PA (CELL-) CELL ACTIVATION INC.
 PA (REGC) UNIV CALIFORNIA.
 PA (SCR1) SCRIPPS RES INST.
 XX
 PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E,
 XX WPI; 1999-580234/49.
 DR

XX Use of cell activating compositions in developing products for
PT diagnosis and treatment of e.g. cardiovascular, inflammatory, rejection,
PT autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection,
PT diabetes, stroke or ischemia -
XX
XX
XX Example 9; Page 182; 184pp; English.
XX This invention describes a novel method for the use and preparation of
XX cell activating compositions which involves preparing a cell activating
XX composition comprising (a) homogenizing pancreatic tissue in buffer at
XX about neutral or higher pH to produce a homogenate; (b) removing
XX particulates from the homogenate; (c) optionally incubating the
XX resulting homogenate with particulates removed, with a protease; and
XX (d) fractionating, homogenate and selecting fractions that exhibit
XX cell activation activity. The methods can be used for improving
XX treatment of inflammatory disease, trauma, autoimmune diseases, arthritis,
XX disease, rejection, diabetes and diabetic complications, stroke, ischemia,
XX Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic
XX retinopathy, diabetes, venous insufficiency, unstable angina or trauma.
XX They can be used in the veterinary treatment of a non-human subject.
XX Protease inhibitors can be used to lower cell activation resulting from
XX these diseases and deficiencies. The detection of an elevated level of
XX hydrogen peroxide can be used to detect an inflammatory condition. An
XX elevated level of hydrogen peroxide in plasma or whole blood and in the
XX presence of superoxide dismutase (SOD) indicates leukocyte up
XX regulation, e.g. indicative of the onset of an acute cardiovascular
XX disorders, such as disease onset or ischemic complications. An elevated
XX level of hydrogen peroxide in plasma or whole blood and a low level in
XX the presence of SOD is indicative of a chronic or immune compromised
XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
XX used in the method of the invention.
XX
XX Sequence 13 AA;
SQ Query Match 100.0%; Score 68; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGVPAIQPVLSGL 13
DB 1 CGVPAIQPVLSGL 13
RESULT 2
AAB31577
ID AAB31577 standard; peptide; 13 AA.
XX AAB31577;
AC AAB31577;
XX 20-APR-2001 (first entry)
DT
XX Amino terminal sequence of bovine trypsin A chain.
XX
XX Fish, serine proteinase; pain; acute inflammation; chronic inflammation;
XX arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;
XX rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
XX systemic lupus erythematosus; phlebitis; tendinitis; rash; psoriasis;
XX acne; eczema; facial seborrheic eczema; foreskin infection;
XX athlete's foot; fissulae infection; ulcer; navel infection; wrinkle;
XX scari; keloid; boil; wart; allergic itch; hemorrhoid; wound;
XX fungal infection; autoimmune disease.
XX
XX Bos sp.
XX WO2000/78332-A2.
PN
XX 28-DEC-2000.
PD
XX 15-JUN-2000; 2000WO-1500005.
PF
XX 18-JUN-1999; 99IS-0005086.
PR

(BJAR/) BJARNASON J B.
Bjarnason JB;
WPI; 2001-091493/10.
Fish serine proteinase, useful as a cosmetic, medicament for treating
eczema, psoriasis, arthritis, and in the manufacture of the medicament
for treating, preventing pathogenic diseases involving receptor
mediated binding -
Disclosure; Page 5; 3pp; English.
The specification describes a fish serine proteinase. The proteinases
are useful as medicaments, for treating and preventing a disease in a
human or an animal such as pain, acute inflammation, chronic arthritis,
inflammation, arthritis, inflamed joints, bursitis, osteoarthritis,
rheumatoid arthritis, juvenile rheumatoid arthritis, sepsidinitis, rash,
fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, eczema,
psoriasis, acne, eczema, facial seborrheic eczema, fissulae infections,
face or neck, foreskin infections, athlete's foot, fissulae infections,
infected topical ulcers, navel infections, hemorrhoids, wounds, wound
keloids, boils, warts and allergic infections and immunological and
autoimmune diseases. They are also useful for removing dead or peeling
skin from otherwise healthy skin and for treating or preventing a
disease in which pathogenesis is caused by bacteria, virus, fungus,
parasite or a protozoan. A receptor mediated binding is involved.
The present sequence represents the amino terminal of bovine trypsin A
chain.
Sequence 13 AA;
SQ Query Match 100.0%; Score 68; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGVPAIQPVLSGL 13
DB 1 CGVPAIQPVLSGL 13
RESULT 3
AAY50207
ID AAY50207 standard; Peptide; 15 AA.
XX AAY50207;
AC AAY50207;
XX 12-JAN-2000 (first entry)
DT
XX Neutrophil-activating pancreatic derived peptide 7.
XX
XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
XX trauma; protease inhibitor; hypertension; sepsis.
XX Bos taurus.
XX WO9946367-A2.
PN
XX 16-SEP-1999.
PD
XX 11-MAR-1999; 99WO-US05247.
PF
XX 11-MAR-1998; 98US-0038894.
PR
XX (CELL-) CELL ACTIVATION INC.
XX (REGC) UNIV CALIFORNIA.
XX (SCRI) SCRIPES RES INST.
XX

XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, pathological pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 11; Page 1081; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytototoxic,
CC neuroprotective, neurotropic, immunomodulatory, relaxant, contractile,
CC gynaecological, cardiant and antinflammatory activity, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition, and the susceptibility to one in a
CC subject. Binding partners to the pancreatic cancer antigens can be used to
CC proteins can be identified pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic cancer, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridization probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 146 AA;
XX
XX Query Match 92.6%; Score 63; DB 21; Length 146;
XX Best Local Similarity 92.3%; Pred. No. 0.0051;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX OY 1 CGVPAIQPVLSGL 13
XX DB 23 CGVPAIHPVLSGL 35
XX
XX
XX RESULT 6
XX ID AAU82738
XX AC AAU82738 standard, Protein; 263 AA.
XX
XX AAU82738;
XX
XX 23-APR-2002 (first entry)
XX
XX Amino acid sequence of novel human protease #37.
XX
XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
XX neuronal-associated disease; metabolic disorder; inflammatory disorder;
XX nervous system disorder; sexual dysfunction; pain; mood disorder;
XX hypertension; psychotic disorder; neurological disorder; dyskinesia;
XX viral infection; human immunodeficiency virus; HIV; non-vital infection;
XX ocular disease; cytostatic; enzyme.
XX
XX Homo sapiens.
XX
XX WO200200860-A2.
XX
XX 03-JAN-2002.
XX
XX 26-JUN-2001; 2001WO-US20171.
XX
XX 26-JUN-2000; 2000US-214047P.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plozman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
XX Charyczak G;
XX

DR WPI; 2002-139913/18.
XX N-PSDB; AAK31780.
XX Nucleic acids encoding novel human proteases, useful for useful for
PT treating diseases and disorders such as cancer, immune-related
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
PT inflammatory disorders -
XX
XX Claim 6; Fig 2M; 313pp; English.
XX
XX The present invention relates to the isolation of novel human
CC proteases, and the nucleic acids encoding them. The sequences of
CC the invention are useful for treating diseases and disorders such as
CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated
CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,
CC mood disorders, attention disorders, cognitive disorders, hypotension,
CC hypertension, psychotic disorders, neurological disorders, dyskinesia,
CC (e.g. Alzheimer's disease, Parkinson's disease) and viral
CC The nucleic acids and polypeptides are also useful for treating viral
CC infections caused by human immunodeficiency virus (HIV), and non-viral
CC infections such as ocular disease (e.g. glaucoma) and macular
CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
CC the invention.
XX
XX Sequence 263 AA;
XX
XX Query Match 92.6%; Score 63; DB 23; Length 263;
XX Best Local Similarity 92.3%; Pred. No. 0.0096;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX OY 1 CGVPAIQPVLSGL 13
XX DB 19 CGVPAIHPVLSGL 31
XX
XX
XX RESULT 7
XX ID AAU50208
XX AC AAU50208;
XX
XX 12-JAN-2000 (first entry)
XX
XX Neutrophil-activating pancreatic derived peptide 8.
XX
XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
XX trauma; protease inhibitor; hypertension; sepsis.
XX
XX Unidentified.
XX
XX WO9946367-A2.
XX
XX 16-SEP-1999.
XX
XX 11-MAR-1999; 99WO-US05247.
XX
XX 11-MAR-1998; 98US-0038894.
XX
XX (CELL-) CELL ACTIVATION INC.
XX (REGC) UNIV CALIFORNIA
XX (SCRI) SCRIPPS RES INST.
XX
XX Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
XX WPI; 1999-580234/49.
XX

PT Use of cell activating compositions in developing products for
PT diagnosis and treatment of e.g. cardiovascular, inflammatory,
PT autoimmune Alzheimer's disease, trauma, arthritis, organ rejection,
PT diabetes, stroke or ischemia -
XX Example 9; Page 182; 184pp; English.
XX
XX This invention describes a novel method for the use and preparation of
XX cell activating compositions which involves preparing a cell activating
XX composition comprising (a) homogenizing pancreatic tissue in buffer at
XX about neutral or higher pH to produce a homogenate; (b) removing
XX particulates from the homogenate; (c) optionally incubating the
XX resulting homogenate, with particulates removed, with a protease; and
XX (d) fractionating the homogenate and selecting fractions that exhibit
XX cell activation activity. The methods can be used for improving
XX treatment outcome or reducing risk of treatment of e.g. cardiovascular
XX disease, inflammatory disease, autoimmune diseases, arthritis,
XX organ rejection, diabetes and diabetic complications, stroke, ischemia,
XX Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic
XX retinopathy, diabetes, venous insufficiency, unstable angina or trauma.
XX They can be used in the veterinary treatment of a non-human subject.
XX Protease inhibitors can be used to lower cell activation resulting from
XX these diseases and deficiencies. The detection of an elevated level of
XX hydrogen peroxide can be used to detect an inflammatory condition. An
XX elevated level of hydrogen peroxide in plasma or whole blood and in the
XX presence of superoxide dismutase (SOD) indicates leukocyte activation
XX disorders, such as disease onset or ischemic complications. An elevated
XX level of hydrogen peroxide in plasma or whole blood and a low level in
XX the presence of SOD is indicative of a chronic or immune compromised
XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
XX used in the method of the invention.
XX Sequence 15 AA;
XX
XX Query Match 91.2%; Score 62; DB 20; Length 15;
XX Best Local Similarity 92.3%; Pred. No. 0.00064;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CGVPAIQPVLSGL 13
XX |||||
XX Db 1 CGVPAIPFVLSGL 13
XX
XX RESULT 8
XX AAY50210
XX ID AAY50210 standard; Peptide; 15 AA.
XX AC AAY50210;
XX XX
XX DT 12-JAN-2000 (first entry)
XX XX
XX DE Neutrophil-activating pancreatic derived peptide 10.
XX XX
XX KW Cell activation; pancreas; treatment; cardiovascular disease; trauma;
XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
XX trauma; protease inhibitor; hypertension; sepsis.
XX XX
XX OS Unidentified.
XX XX
XX PN W09946367-A2.
XX XX
XX XX
XX PD 16-SEP-1999.
XX XX
XX PF 11-MAR-1999; 99MC-US05247.
XX XX
XX PR 11-MAR-1998; 98US-0018894.
XX XX
XX PA (CELL-) CELL ACTIVATION INC.
XX (REGC) UNIV CALIFORNIA.
XX PA (SCRI) SCRIPPS RES INST.

XX Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
XX WPI; 1999-580234/49.
XX
XX Use of cell activating compositions in developing products for
XX diagnosis and treatment of e.g. cardiovascular, inflammatory,
XX autoimmune Alzheimer's disease, trauma, arthritis, organ rejection,
XX diabetes, stroke or ischemia -
XX Example 9; Page 182; 184pp; English.
XX
XX This invention describes a novel method for the use and preparation of
XX cell activating compositions which involves preparing a cell activating
XX composition comprising (a) homogenizing pancreatic tissue in buffer at
XX about neutral or higher pH to produce a homogenate; (b) removing
XX particulates from the homogenate; (c) optionally incubating the
XX resulting homogenate, with particulates removed, with a protease; and
XX (d) fractionating the homogenate and selecting fractions that exhibit
XX cell activation activity. The methods can be used for improving
XX treatment outcome or reducing risk of treatment of e.g. cardiovascular
XX disease, inflammatory disease, autoimmune diseases, arthritis,
XX organ rejection, diabetes and diabetic complications, stroke, ischemia,
XX Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic
XX retinopathy, diabetes, venous insufficiency, unstable angina or trauma.
XX They can be used in the veterinary treatment of a non-human subject.
XX Protease inhibitors can be used to lower cell activation resulting from
XX these diseases and deficiencies. The detection of an elevated level of
XX hydrogen peroxide can be used to detect an inflammatory condition. An
XX elevated level of hydrogen peroxide in plasma or whole blood and in the
XX presence of superoxide dismutase (SOD) indicates leukocyte up
XX regulation, e.g. indicative of the onset of an acute cardiovascular
XX disorder, such as disease onset or ischemic complications. An elevated
XX level of hydrogen peroxide in plasma or whole blood and a low level in
XX the presence of SOD is indicative of a chronic or immune compromised
XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
XX used in the method of the invention.
XX Sequence 15 AA;
XX
XX Query Match 91.2%; Score 62; DB 20; Length 15;
XX Best Local Similarity 92.3%; Pred. No. 0.00064;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CGVPAIQPVLSGL 13
XX |||||
XX Db 1 CGVPAIPFVLSGL 13
XX
XX RESULT 9
XX AAB54077
XX ID AAB54077 standard; Protein; 192 AA.
XX AC AAB54077;
XX XX
XX DT 09-MAR-2001 (first entry)
XX XX
XX DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.
XX XX
XX KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
XX detection; diagnosis; identification; cytostatic; neuroprotective;
XX antineoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
XX antifibrotic; radiant; gene therapy; chromosome mapping;
XX linkage analysis; tissue identification; tissue typing; forensic;
XX neural; immune system; vascular; reproductive; gastrointestinal;
XX pulmonary; cardiovascular; renal; proliferative.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200055320-A1.
XX XX
XX XX 21-SEP-2000.
XX XX

08-NAR-2000; 2000WO-US05989.	XX
12-NAR-1999; 99US-0124270.	XX
(HUMA-) HUMAN GENOME SCI INC.	XX
Rosen CA, Ruben SM;	XX
WPI; 2000-579444/54.	XX
N-PSDB; AAC98842.	XX
New nucleic acid that is a pancreatic cancer antigen for preventing, treating, or ameliorating a medical condition, particular pancreatic cancer, or for use in assays for diagnosing a pathological condition -	XX
Claim 11; Page 966; 1379pp; English.	XX
AAC98773 to AAC99231 encode the human pancreatic cancer associated proteins, called pancreatic cancer antigens, given in AAB54008 to AAB54466. The human pancreatic cancer antigens have cytostatic, neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive, gynaecological, cardiac and antiinflammatory activities, and can be used in gene therapy. The polynucleotide and proteins can be used for preventing, treating, or ameliorating a medical condition or in assays for diagnosing a pathological condition or a susceptibility to one in a subject. Binding partners to the proteins and the activity of the proteins can be identified. The pancreatic cancer antigens can be used to detect, treat or prevent pancreatic disorders, especially cancer. Agonists and antagonists to the antigens can be screened for. The pancreatic cancer antigen polynucleotides can be used to design nucleic acid hybridization probes and/or typing and a variety of forensic analysis, tissue identification and a variety of forensic and diagnostic methods. The proteins can be used to generate antisera and diagnostic methods, detect and target the polypeptides, including which are used in vitro diagnostic and therapeutic methods. The proteins can be used to treat or prevent neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal or proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent sequences used in the exemplification of the present invention.	XX
Sequence 192 AA;	XX
Sequence 192 AA;	XX

Query Match	72.1%	Score 49;	DB 21;	Length 192;
Best Local Similarity	72.7%	Pred. No. 1.7;		
Mismatches	8;	Conservative 2;	Mismatches 1;	Indels 0;
Gaps	0;			

QY 1 CGVPAIQPVLS 11
|||:|:|:|
nb 15 CGIPAIKPALS 25

RESULT 10
AAB11710
IN AAB11710 standard: Protein; 264 AA.

XX	
AC	AAB11710;
XX	22 DEC 2000 (first entry)

[illegible]

XX BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
 KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
 KW epilepsy; cancer; inflammation; infertility; pancreatitis;
 KW prostatic hypertrophy.

XX
05
Homo sapiens.

XX
PN
W0200031243-A1.

XX 03-TTN-2000

XX
10 NOV-1999. 99W0-TP06473.

10
9
8
7

XX 20-NOV-1998; 98JP-0347806.
PR (FUSO) FUSO PHARM IND LTD.
PA Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsu S:
PJ WPI : 2000-400058/34.
 PP N-PSTDB; AAA61733 .
RR Serine proteases BSPPs, useful in detecting homologs, mutants and
SX polymorphic variants as markers for diagnosis of e.g. Alzheimer's
TU diseases, epilepsy, cancer and inflammation, using blood, urine,
VZ pancreas or other tissues -
XX Claim 1. Page 51-52; 70pp.; Japanese.

The invention relates to novel serine proteases designated BSSP5 (AAB1170-B1171), and to nucleic acids encoding these (AAB6173-A61734). The invention also relates to vectors and transgenic animals comprising BSSP5 nucleic acids; transgenic animals in which the expression level of BSSP5 can be varied; and a BSSP5 knockout method. The invention additionally encompasses anti-BSSP5 antibodies and methods of production of such antibodies, methods of BSSP5 detection using the antibodies, and the use of BSSP5 proteins or fragments. A method for detecting medical conditions, e.g., pancreatitis. A method for detecting pancreatitis comprising measuring BSSP5 concentration in the blood or urine, and a pancreatic nucleotides agent containing an anti-BSSP5 antibody is also disclosed. Nucleotides encoding BSSP5 were initially isolated from a human brain cDNA library using degenerate PCR primers (AAB6174-A61745) based on conserved regions of serine proteases. The BSSP5 serine proteases and nucleotides encoding them are useful in detecting hemoglobins, mutants and polymorphic variants in biological samples (e.g., blood, urine, brain, prostate gland, placenta, testis, pancreas and spleen) as diagnostic markers for conditions such as Alzheimer's disease, epilepsy, cancer, inflammation, infertility, pancreatitis and prostatic hypertrophy. Sequence AAB1170 represents human BSSP5 (hBSSP5), and sequence AAB1171 represents murine BSSP5 (mBSSP5).

XX
CO
commence 264 AA:

Query March 72.1%; Score 49; DB 21; Length 264;
Best Local Similarity 72.7%; Pred. No. 2.4;
Matches 8; Conservative 2; Mismatches 1; Indels

QY 1 CGVPAIQPVLS 11
||:||||:|
nB 19 CGIPAIPKPLS 29

```

RESULT 11
AAB11711
ID AAB11711 standard; Protein; 264 AA.
XX
AC AAB11711;
XX
DT 23-OCT-2000 (first entry)
XX

```

XX BSS5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
KW epilepsy; cancer; inflammation; infertility; pancreatitis;
KW prostatic hypertrophy.

50
YY
MUG BP.

XX
PN
WO200031243-A1.

XX 02-1111-2000

XX 1000.

```

XX 20-NOV-1998; 98JP-0347806.
XX (FUSO) FUSO PHARM IND LTD.
XX
XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
XX MPI: 2000-400058/34.
XX N-PSDB; AAA61734.
XX
XX Serine proteases BSSPs, useful in detecting homologs, mutants and
XX polymorphic variants as markers for diagnosis of e.g. Alzheimer's
XX disease, epilepsy, cancer and inflammation, using blood, urine,
XX pancreas or other tissues
XX
XX Claim 3; Page 55-56; 70pp; Japanese.
XX
XX The invention relates to novel serine proteases designated BSSPs
XX (AAB1170, AAB1171, and to nucleic acids encoding them (AAA61733-A61734).
XX The invention also relates to vectors and transformants comprising BSSPs
XX nucleic acids, and transgenic animals in which the expression level of BSSPs
XX can be varied; and transgenic knockout mouse. The invention additionally
XX encompasses anti-BSSPs antibodies and methods of production of such
XX antibodies, methods of BSSPs detection, and methods of using the
XX use of BSSPs proteins or fragments as diagnostic markers for certain
XX medical conditions, e.g., pancreatitis. A method for detecting
XX pancreatitis comprising measuring BSSPs concentration in blood or
XX urine, and a pancreatitis diagnostic agent containing an anti-BSSPs
XX antibody is also disclosed. Nucleotides encoding BSSPs were initially
XX isolated in a human brain cDNA library using degenerate PCR primers
XX (AAB61744-A61745) based on conserved regions of serine proteases. The
XX BSSPs serine proteases and nucleotides encoding them are useful in
XX detecting homologues, mutants and polymorphic variants in biological
XX samples, e.g., blood, urine, brain, prostate gland, placenta, testis,
XX pancreas (epitum) as diagnostic markers for conditions such as
XX Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
XX pancreatitis and prostatic hypertrophy. Sequence AAB1170 represents
XX human BSSPs (hBSSPs), and sequence AAB1171 represents murine BSSPs
XX (mBSSPs).
XX
XX Sequence 264 AA;
XX
XX Query Match 70.6%; Score 48; DB 21; Length 264;
XX Best Local Similarity 81.8%; Pred. No. 3.6;
XX Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CGVPAIQPVLS 11
XX |||||
XX Db 19 CGVPAITPAIS 29
XX
XX RESULT 12
XX AAY50212
XX ID AAY50212 standard; Peptide, 13 AA.
XX AC AAY50212;
XX
XX DT 12-JAN-2000 (first entry)
XX DE Neutrophil-activating pancreatic derived peptide 12.
XX
XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
XX trauma; protease inhibitor; hypertension; sepsis.
XX
XX OS Unidentified.
XX
XX WO9946367-A2.
XX
XX 16-SEP-1999.
XX

```

```

PF 11-MAR-1999; 99WC-US05247.
XX
XX 11-MAR-1998; 98US-0038894.
XX
XX (CELL-) CELL ACTIVATION INC.
XX (REGC) UNIV CALIFORNIA
XX (SCRI) SCRIPPS RES INST.
XX
XX Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;
XX WPI: 1999-580234/49.
XX
XX Use of cell activating compositions in developing products for
XX diagnosis and treatment of e.g. cardiovascular, inflammatory,
XX autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection,
XX diabetes, stroke or ischemia
XX
XX Example 9; Page 182; 184pp; English.
XX
XX This invention describes a novel method for the use and preparation of
XX cell activating compositions which involves preparing a cell activating
XX composition comprising (a) homogenizing pancreatic tissue in buffer at
XX about neutral or higher pH to produce a homogenate; (b) removing
XX particulates from the homogenate; (c) optionally incubating the
XX resulting homogenate, with particulates removed, with a protease, and
XX (d) fractionating the homogenate and selecting fractions that exhibit
XX cell activation activity. The methods can be used for improving
XX treatment outcome or reducing risk of treatment of e.g. cardiovascular
XX disease, inflammatory disease, trauma, autoimmune diseases, arthritis,
XX organ rejection, diabetes and diabetic complications, stroke, ischemia,
XX Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic
XX neuropathy, diabetes, venous insufficiency, unstable angina or trauma.
XX These diseases and deficiencies. The detection of an elevated level of
XX protease inhibitor in the veterinary treatment of a non-human subject.
XX Protease inhibitor can be used. The detection of an elevated level of
XX hydrogen peroxide can be used. The detection of an inflammatory condition. An
XX elevated level of hydrogen peroxide in plasma or whole blood and in the
XX presence of superoxide dismutase (SOD) indicates acute cardiac
XX regulation, e.g. indicative of the onset of an acute cardiac
XX disorders, such as disease onset or ischemic complications. An elevated
XX level of hydrogen peroxide in plasma or whole blood and a low level in
XX the presence of SOD is indicative of a chronic or immune compromised
XX condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides
XX used in the method of the invention.
XX
XX Sequence 13 AA;
XX
XX Query Match 67.6%; Score 46; DB 20; Length 13;
XX Best Local Similarity 80.0%; Pred. No. 0.31;
XX Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 CGVPAIQPVL 10
XX |||||
XX Db 1 CGVPAIKPAL 10
XX
XX RESULT 13
XX AAY50211
XX ID AAY50211 standard; Peptide; 11 AA.
XX AC AAY50211;
XX
XX DT 12-JAN-2000 (first entry)
XX DE Neutrophil-activating pancreatic derived peptide 11.
XX
XX Cell activation; pancreas; treatment; cardiovascular disease; trauma;
XX inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke;
XX organ rejection; ischemia; Alzheimer's disease; myocardial infarction;
XX haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina;
XX trauma; protease inhibitor; hypertension; sepsis.
XX
XX OS Unidentified.
XX

```

	Homo sapiens.
XO	
XX	WO200175067-AZ.
PX	
NN	
PP	
DD	11-OCT-2001..
XX	
XX	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
PA	(HYSE-) HYSEP INC.
XX	
XX	Dzmanac RT, Liu C, Tang YT;
XI	
WI	MPI: 2001-639362/73.
DR	N-PSTDB; AAS76824.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
Bio-	identity -
XXXX	
Claim	20; SEQ ID No 44996; 103pp: English.
XX	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, detection of, for chromosome
CC	mapping and gene mapping, and used in diagnostics of, (III). The
CC	polyclonally expressed genes. (I) is useful in gene therapy techniques
CC	for identifying expressed genes. (I) is useful in disease states involving
CC	to restore normal activity of (II) or to treat disease states involving or
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight marker, as a
CC	food supplement. (II) and its binding partners are useful for treating
CC	imaging of sites expressing (II), and its expression or biological activity.
CC	disorders involving abnormality of sequence of the invention such as
CC	polymorphic diseases, forensic science, gene mapping have applications in
CC	diseases, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABQ0010-ABG0307 represent novel human
CC	diagnostic amino acid sequences of the invention which appear in the printed
CC	Note: The sequence data for each protein has been prepared according to
CC	specification, for publication in electronic format directly from WIPO
at	ftp://wipo.int/pub/published_pct_sequences.
Sequence	107 AA;
Query Match	61.8%; Score 42; DB 22; Length 107;
Best Local Similarity	66.7%; Pres.No.: 15;
Score	2; Indexes:
Gaps	0;

DB	11 GDPQLQPVLAGL 22
RESULT 15	
AAU68565	
ID	AAU68565 standard; Protein: 107 AA.
XX	
AC	AAU68565;
XX	
DT	16-JAN-2002 (first entry)
XX	
DE	Human novel cytokine encoded by cDNA 790CIP2A_11 #2.
XX	
XX	Human; cytokine; cell proliferation; cell differentiation;
KW	transformin; transforming growth factor; stem cell growth factor; activin; inhibin;
XX	nervous system disease; neuropathy; Alzheimer's disease;
KW	Parkinson's disease; Huntington's disease; spinal cord disorder;

KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW osteitic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoarthritis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.

XX Homo sapiens.

OS
 XX WO200175093-A1.

XX
 XX 11-OCT-2001.

XX
 XX 30-MAR-2001; 2001WO-US10484.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0643147.

PR 22-SEP-2000; 2000US-0668680.

PR 23-OCT-2000; 2000US-0695618.

PR 30-NOV-2000; 2000US-0728711.

PR 14-MAR-2001; 2000US-0728711.

XX (HYSE-) HYSEQ INC.

XX
 XX Tang Yr, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;

PI Yang Y, Zaho Qa, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;

XX WPI; 2001-626432/72.

DR N-PSDB; AAS59857.

XX
 XX New polypeptides and nucleic acids, useful for diagnosis, treatment of

PT inflammatory autoimmune, neurological, myeloid or lymphoid cell, bone

PT degenerative disorders, cancer and promoting wound healing

XX Claim 20; Page 308; 336pp; English.

XX
 XX The invention relates to isolated human polypeptides (which may be

CC cytokines) and the polynucleotides encoding them. The proteins are

CC for identifying a compound which binds to it (e.g. modulators

CC and antagonists). The polynucleotides are useful as an array for

CC detection. The proteins and nucleic acids are useful as nutritional

CC sources or supplements. The protein exhibits activity relating

CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,

CC stem cell growth factor activity, immune stimulating or immune

CC inhibiting and activin or inhibin related activities. The proteins (and

CC antibodies raised against them) and nucleic acids are therefore useful in

CC the diagnosis and treatment of diseases and disorders such as cancer,

CC central and peripheral nervous system diseases and neuropathies.

CC Alzheimer's, Parkinson's, Huntington's disease, amyotrophic

CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular

CC diseases, stroke, myeloid or lymphoid disorders, platelet disorders,

CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for

CC regeneration of bone, cartilage, tendon, ligament and connective tissue

CC growth, and in tissue repair, healing of burns, incision, lacerations, or

CC treating osteoporosis, osteoarthritis, bone degenerative disorder, or

CC periodontal disease, lung or liver fibrosis, reperfusion injury in

CC various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
 CC such as haemophilia, septic shock, sepsis, arthritis, nephritis and inflammatory
 CC bowel disease. Virus infection and are useful in altering bodily
 CC characteristics. The present sequence represents a novel protein of the
 CC invention.

XX Sequence 107 AA;

Query Match 61.8%; Score 42; DB 22; Length 107;

Best Local Similarity 66.7%; Pred. No. 15;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 GYPAIQPVLSQL 13

Db 11 GDPELQVPLVAGL 22

Search completed: February 12, 2003, 10:22:25

Job time : 9.73134 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50, Search time 3.2597 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-5

Perfect score: 68

Sequence: 1 CGVPAIQPVLSGL 13

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	245	1 KYBOA	chymotrypsin (EC 3
2	68	100.0	245	1 KYBOA	chymotrypsin (EC 3
3	68	100.0	263	2 A21195	chymotrypsin (EC 3
4	63	92.6	263	2 A31299	chymotrypsin (EC 3
5	62	91.2	16	2 C61414	chymotrypsin (EC 3
6	62	91.2	17	2 B61414	chymotrypsin (EC 3
7	61	89.7	263	1 KYR7B	chymotrypsin (EC 3
8	58	85.3	28	2 A61529	chymotrypsin (EC 3
9	54	79.4	20	2 A61414	chymotrypsin (EC 3
10	53	77.9	25	2 A45109	chymotrypsin (EC 3
11	43	72.1	264	2 I38136	chymotrypsin (EC 3
12	47	62.1	163	2 S47537	chymotrypsin-like
13	46	67.6	246	2 A23473	chymotrypsin-like
14	46	67.6	246	2 A23473	chymotrypsin-like
15	44.5	65.4	384	2 A30219	chymotrypsin B - A
16	41	60.3	342	2 A33450	ceramide glucosylt
17	40	58.8	255	2 B87595	proline racemase (
18	40	58.8	475	1 RKWL	transcription regu
19	39.5	58.1	364	2 A44899	ribulose-bisphosph
20	39.5	58.1	396	1 TRYXBA	ribulose-bisphosph
21	39	57.4	86	2 AC2549	ribulose-bisphosph
22	39	57.4	241	2 C81971	chymotrypsin-like
23	39	57.4	241	2 H81026	hypothetical prote
24	39	57.4	282	2 A84341	conserved hypotab
25	39	57.4	523	1 A44195	hypothetical prote
26	39	57.4	910	2 A10796	UL21 protein homol
27	38	55.9	111	2 C71227	NAOH2 dehydrogen
28	38	55.9	206	2 H47021	pectic enzyme secr
29	38	55.9	268	2 S68825	pancreatic elastas

30 38 55.9 268 2 S68825
31 38 55.9 269 2 C70710
32 38 55.9 400 2 A10261
33 38 55.9 452 2 C82099
34 38 55.9 452 2 C82099
35 38 55.9 553 2 T03825
36 38 55.9 829 2 T19484
37 38 55.9 1048 1 XP8E9A
38 37.5 55.1 226 2 T35172
39 37.5 55.1 237 2 A70672
40 37.5 55.1 345 2 T32416
41 37.5 55.1 466 1 TMFF
42 37.5 55.1 525 2 T25550
43 37.5 55.1 557 2 S21596
44 37 54.4 126 2 G84399
45 37 54.4 378 2 T19012

ALIGNMENTS

RESULT 1
KYBOA
chymotrypsin (EC 3.4.21.1) B precursor - bovine
N/Alternate names: chymotrypsinogen B
C/Species: Bos primigenius taurus (cattle)
C/Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 18-Jul-1997
C/Accession: A00953
R/Smillie, L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.
Nucleotide position: 218, 343-346, 1968
A/Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen A
A/Reference number: A00953; MUID:68238908; PMID:5649671
A/Accession: A00953
A/Molecule type: protein
A/Residues: 1-245 <SKI>
C/Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acini of the pancreas. The first activation cleavage, leading to pi-chymotrypsin B, occurs in the sar
C/Keywords: hydrolytic; enzyme; protein digestion; serine proteinase; zymogen
F/1-15/Domain: propetide; zymogen; experimental <PRO>
F/16-245/Product: chymotrypsin B #status experimental <MAT>
F/1-122,42-58,136-201,168-182,191-220/Diulfide bonds: #status experimental
F/57,102,195/Active site: His, Asp, Ser #status experimental

Query Match
Best Local Similarity 100.0%; Score 68; DB 1; Length 245;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSGL 13

Db 1 CGVPAIQPVLSGL 13

RESULT 2

KYBOA

chymotrypsin (EC 3.4.21.1) A precursor - bovine

N/Alternate names: chymotrypsinogen A

C/Species: Bos primigenius taurus (cattle)

C/Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999

C/Accession: A90235; A93158; S29650; A00952

R/Brown, J.R.; Hartley, B.S.

Biochem. J. 101, 214-228, 1966

A/Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphid

A/Reference number: A90235; MUID:67181721; PMID:5971783

A/Accession: A90235

A/Molecule type: protein

A/Residues: 1-101, N', 101-245 <BRO>

R/Brown, J.R.; Birktoft, J.J.; Hartley, B.S.

Nucleotide position: 221, 337-340, 1969

A/Title: Role of a buried acid group in the mechanism of action of chymotrypsin.

A/Reference number: A93158; MUID:6506266; PMID:5764436

A/Contents: annotation; revision to residue 102

R;Maloun, B.; Klueh, I.; Kostka, V.; Moravsek, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Soml
 Biochim. Biophys. Acta 130, 543-546, 1966
 A;Title: Covalent structure of bovine chymotrypsinogen A.
 A;Reference number: A90572; MUID:67103948; PMID:5972866
 A;Accession: A93158
 A;Molecule type: protein
 A;Residues: 1-101, 'N', 103-245 <MEL>
 A;Note: disulfide bonds were determined
 R;Cutruzzola, F.; Ascenzi, P.; Barra, D.; Bolognesi, M.; Menegatti, E.; Sarti, P.; Schne
 Biochim. Biophys. Acta 1161, 201-208, 1993
 A;Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catalytic
 A;Reference number: S29650; MUID:93160238; PMID:8431470
 A;Accession: S29650
 A;Molecule type: protein
 A;Residues: 1-12;16-27;149-160;181-200 <CUT>
 R;Smillie, L.B.; Harclay, B.S.
 Biochem. J. 101, 232-241, 1966
 A;Title: Histiidine sequences in the active centres of some 'serine' proteinases.
 A;Reference number: A90230; MUID:67081733; PMID:5971785
 A;Content: annotation; site
 R;Fairbrother, J.G.; Bloch, D.M.; Henderson, R.; Steltz, T.A.
 J. Mol. Biol. 100, 257-267, 1976
 A;Title: The structure of alpha-chymotrypsin.
 A;Reference number: A93754
 A;Content: annotation; X-ray crystallography
 C;Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.
 C;Comment: Trypsin cleavage after Arg-15 results in a fully active enzyme (pi-chymotrypsin
 delta-chymotrypsin; further chymotryptic cleavage liberates the dipeptide Thr-147 and Arg-
 d Arg-148 directly from chymotrypsinogen, which leads to the degraded form neo-chymotrypsin
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F;1-1245/Product: chymotrypsinogen #status experimental <ZYMS>
 F;1-13,16-146,149-245/Product: alpha-chymotrypsin #status experimental <MPT>
 F;16-238/Domain: trypsin homology <TRY>
 F;1-122,42-58,136-201,168-182,191-220/Dsulfide bonds: #status experimental
 F;57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 100.0%; Score 68; DB 1; Length 245;
 Best Local Similarity 100.0%; Pred. No. 0.00057; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0

QY 1 CGVPAIQPVLSGL 13
 DB 1 CGVPAIQPVLSGL 13

RESULT 3
 A21195
 Chymotrypsin (EC 3.4.21.1) 2 precursor - dog
 C;Species: Canis lupus familiaris (dog)
 C;Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999
 C;Accession: A21195
 R;Pinkey, S.D.; Leforge, K.S.; Luc, V.; Scheele, G.
 Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983
 A;Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence data
 A;Reference number: A21195; MUID:84170253; PMID:6584866
 A;Status: preliminary
 A;Accession: A21195
 A;Molecule type: mRNA
 A;Residues: 1-263 <PIN>
 A;Cross-references: GB:K01173; NID:G163945; PIDN:AAA30841.1; PID:G163946
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; protein digestion; serine proteinase
 F;34-256/Domain: trypsin homology <TRY>
 F;75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 68; DB 2; Length 263;
 Best Local Similarity 100.0%; Pred. No. 0.00061; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0

QY 1 CGVPAIQPVLSGL 13
 DB 19 CGVPAIQPVLSGL 31

RESULT 4

A31299

Chymotrypsin (EC 3.4.21.1) precursor - human

C;Species: Homo sapiens (man)
 C;Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 22-Jun-1999

C;Accession: A31299

R;Tonita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsumi
 Biochem. Biophys. Res. Commun. 158, 569-575, 1989A;Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinogen
 A;Reference number: A31299; MUID:89134264; PMID:2917002

A;Accession: A31299

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-263 <TON>

A;Cross-references: GB:M24400; NID:G181189; PIDN:AAA52128.1; PID:G181190

C;Genetics: GDB:CTRB1; CTRB

A;Cross-references: GDB:119820; OMIM:118890

A;Map position: 16q23.1-16q23.1

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; protein digestion; serine proteinase

F;34-256/Domain: trypsin homology <TRY>

F;75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 92.6%; Score 63; DB 2; Length 263;

Best Local Similarity 92.3%; Pred. No. 0.0043;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13

DB 19 CGVPAIQPVLSGL 31

RESULT 5

C61414

Chymotrypsin (EC 3.4.21.1) - slider turtle (fragment)

C;Species: Pseudemys scripta (slider)

C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999

C;Accession: C61414

R;Bhatnagar, A.K.; Barnard, E.A.

J. Mol. Evol. 2, 187-198, 1973

A;Title: Evolution in the pancreatic chymotrypsinogen series: N-terminal sequence determ

A;Reference number: A61414; MUID:76146602; PMID:4807189

A;Accession: C61414

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-16 <BHA>

C;Keywords: hydrolase; protein digestion; serine proteinase

Query Match 91.2%; Score 62; DB 2; Length 16;

Best Local Similarity 92.3%; Pred. No. 0.00042;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13

DB 1 CGVPAIQPVLSGL 13

RESULT 6

B61414

Chymotrypsin (EC 3.4.21.1) - painted turtle (fragment)

C;Species: Chrysemys picta (painted turtle)

C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 07-May-1999

C;Accession: B61414

R;Bhatnagar, A.K.; Barnard, E.A.

J. Mol. Evol. 2, 187-198, 1973

A;Title: Evolution in the pancreatic chymotrypsinogen series: N-terminal sequence determ

A;Reference number: A61414; MUID:76146602; PMID:4807189

A;Accession: B61414

A;Status: preliminary

A;Molecule type: protein

F;75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 91.2%; Score 62; DB 2; Length 16;

Best Local Similarity 92.3%; Pred. No. 0.00042;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13

DB 1 CGVPAIQPVLSGL 13


```

A:Residues: 1-17 <BHA>
C:Keywords: hydrolase; serine proteinase

Query Match      91.2%; Score 62; DB 2; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.00044;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
Db 1 CGVPAIPVLSGL 13

RESULT 7
KYRTB
Chymotrypsin (EC 3.4.21.1) B precursor - rat
C:Alternate names: chymotrypsinogen B
C:Species: Rattus norvegicus (Norway rat)
C:Accession: A61529 #sequence_revision 28-Dec-1997 #text_change 18-Jun-1999
R:Bell G.V., Osoegawa, C.; Ooiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.
J. Biol. Chem. 255: 14265-14270, 1980
A:Title: Isolation and sequencing of rat chymotrypsin B gene.
A:Reference number: A22658; MUID:85054881; PMID:6209274
A:Accession: A22658
A:Reference number: A22658
A:Molecule type: DNA
A:Residues: 1-263 <BEL>
A:Cross-references: GB:K02298; NID:g203653; PIDN:AA98732.1; PID:g203654
C:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:19-33/Domain: propeptide #status predicted <PRO>
F:34-252/Product: chymotrypsin B #status predicted <MAT>
F:75-120,213/Active site: His, Asp, Ser #status predicted

Query Match      89.7%; Score 61; DB 1; Length 263;
Best Local Similarity 84.6%; Pred. No. 0.0093;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
Db 19 CGVPTIQPVLTGL 31

RESULT 8
A61529
Chymotrypsin (EC 3.4.21.1) - Atlantic cod (fragments)
C:Species: Gadus morhua (Atlantic cod)
C:Accession: A61529 #sequence_revision 15-Oct-1994 #text_change 11-May-2000
R:Asgeirsson, B.; Bjarnason, J.B.
Comp. Biochem. Physiol. B 99: 397-395, 1991
A:Title: Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua)
A:Reference number: A61529; MUID:92111252; PMID:1764912
A:Accession: A61529
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-28 <ASG>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match      85.3%; Score 58; DB 2; Length 28;
Best Local Similarity 76.9%; Pred. No. 0.0034;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
Db 1 CGSPAIQPVLSGI 13

RESULT 9
A61529
Chymotrypsin (EC 3.4.21.1) - snapping turtle (fragment)
C:Species: Chelydra serpentina (snapping turtle)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
C:Accession: A61414
R:Bhargava, A.K.; Barnard, E.A.
J. Mol. Evol. 2: 187-198, 1973
A:Title: Evolution in the pancreatic chymotrypsinogen series: N-terminal sequence determination
A:Reference number: A61414; MUID:76146602; PMID:4807189
A:Accession: A61414
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <BHA>
C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match      79.4%; Score 54; DB 2; Length 20;
Best Local Similarity 84.6%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
Db 1 CWPPIQPVLSGL 13

RESULT 10
A45109
Chymotrypsin (EC 3.4.21.1) homolog p31 - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 31-Mar-1997
C:Accession: A45109
R:Brill, S.U.; Hamilton, J.W.; Cohn, D.V.
J. Biol. Chem. 267: 21595-21600, 1992
A:Title: Roentgen crystallographic study of porcine chymotrypsin: a non-constitutive, secretory protein binds porcine chymotrypsin
A:Reference number: A45109; MUID:93016107; PMID:1400470
A:Accession: A45109
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GOR>
A:Experimental source: pancreas
A>Note: sequence extracted from NCBI backbone (NCBI:P:116723)
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match      77.9%; Score 53; DB 2; Length 25;
Best Local Similarity 91.7%; Pred. No. 0.021;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVPPIQPVLSGL 13
Db 2 GVPPIPVVLSGL 13

RESULT 11
138136
Chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
C:Accession: I38136
R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.
Hum. Mol. Genet. 2: 1589-1595, 1993
A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.
A:Reference number: I38135; MUID:94093544; PMID:8268911
A:Accession: I38136
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <RES>
A:Cross-references: ENBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228
C:Genetics:
A:Gene: GDB:CTRL
A:Cross-references: GDB:204061
A:Map position: 16q22.1-16q22.1
A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3
C:Superfamily: trypsin; trypsin homology

```

C;Keywords: hydrolase; serine proteinase
 F:34-257/Domain: trypsin homology <TRY>
 F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 72.1%; Score 49; DB 2; Length 264;
 Best Local Similarity 72.7%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 11
 Db 19 CGVPAIKPALS 29

RESULT 12
 547537
 chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
 C;Species: Gadus morhua (Atlantic cod)
 C;Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C;Accession: S47537; S43163
 R;Gudmundsdottir, A.; Olgason, S.; Eskin, A.E.; Craik, C.S.; Bjarnason, J.B.
 Biochim. Biophys. Acta 1219, 211-214, 1993
 A;Title: Atlantic cod cDNA encoding a serine chymotrophic chymotrypsinogen.
 A;Reference number: S47537; MUID:94368860; PMID:8086467
 A;Accession: S47537A
 A;Molecule type: mRNA
 A;Residues: 1-263 <CD>
 C;Superfamily: trypsin; trypsin homology
 C;Cross-references: KMBL:X78490; NID:9468750; PIDN:CAA55242.1; PID:9468751
 A;Title: chymotrypsin; trypsin homology; serine proteinase
 F:19-263/Domain: signal sequence #status predicted <SIG>
 F:19-263/Product: chymotrypsin #status predicted <MAT>
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 69.1%; Score 47; DB 2; Length 263;
 Best Local Similarity 66.7%; Pred. No. 2.2;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 12
 Db 19 CGRPAISPVTG 30

RESULT 13
 A23473
 chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)
 N;Alternate names: pancreatic elastase II (misidentification)
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1997 #text_change 31-Mar-2000
 C;Accession: A23473
 R;Vered, M.; Gertler, A.; Burstein, Y.
 Int. J. Pept. Protein Res. 27, 183-190, 1986
 A;Reference number: A23473; MUID:86194934; PMID:3634756
 A;Accession: A23473
 A;Molecule type: protein
 A;Residues: 1-126 <VER>
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; serine proteinase

Query Match 67.6%; Score 46; DB 2; Length 126;
 Best Local Similarity 80.0%; Pred. No. 1.6;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 10
 Db 1 CGVPAIKPALS 10

RESULT 14
 S72219
 chymotrypsin B - Atlantic cod (fragments)
 C;Species: Gadus morhua (Atlantic cod)
 C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Aug-1998

C;Accession: S72219
 R;Eich-Larsen, R.; Ageirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.
 Biochim. Biophys. Acta 1297, 49-56, 1996
 A;Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.
 F:75,121,214/Active site: His, Asp, Ser #status predicted
 A;Reference number: S72219; MUID:96439045; PMID:8841380
 A;Accession: S72219
 A;Status: Preliminary
 A;Molecule type: protein
 A;Residues: 1-14,15-244 <LET>
 C;Superfamily: trypsin; trypsin homology
 F:15-237/Domain: trypsin homology <TRY>

Query Match 67.6%; Score 46; DB 2; Length 244;
 Best Local Similarity 66.7%; Pred. No. 3;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 12
 Db 1 CGSPAIPQVGTG 12

RESULT 15
 AH0241
 ceramide glucosyltransferase (EC 2.4.1.80) [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AH0241
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I.
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AH0001; MUID:21470413; PMID:11586360
 A;Accession: AH0241
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-384 <KOR>
 A;Cross-references: GB:AL590842; PIDN:CAC90796.1; PID:gi5979996; GSPDB:GN00175
 C;Genetic81
 A;Gene: IPO1993
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 65.4%; Score 44.5; DB 2; Length 384;
 Best Local Similarity 58.8%; Pred. No. 8.2;
 Matches 10; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 1 CGVPA---AIQPVLSG 12
 Db 31 CGVPAQATILQPVLSG 47

Search completed: February 12, 2003, 10:28:51
 Job time : 5.2597 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:04:45 ; Search time 1.74627 Seconds
(with alignment)
308.768 Million cell updates/sec

Title: US-10-036-371-5

Perfect score: 68

Sequence: 1 CGVPAIQPVLSGL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	245	1 CTRA_BOVIN	P00766 bos taurus
2	68	100.0	245	1 CTRB_BOVIN	P00767 bos taurus
3	68	100.0	263	1 CTR2_CANFA	P04813 canis famil
4	61	82.6	263	1 CTRB_HUMAN	P17538 homo sapien
5	61	82.6	263	1 CTRB_RAT	P07338 rattus norv
6	49	72.1	263	1 CTRB_HUMAN	P40313 homo sapien
7	47	69.1	263	1 CTRB_HUMAN	P47796 gadus morhu
8	46	67.6	245	1 CTRB_HUMAN	P80646 gadiu morhu
9	40	58.8	475	1 RBL_CHLRE	P08877 chlamydomon
10	40	58.8	591	1 HE_HEMPT	P1953 chlamydomon
11	39.5	58.1	364	1 Y084_CAEEL	P34620 caenorhabdi
12	39.5	58.1	397	1 PRLA_LYSEN	P08077 chlamydomon
13	39.5	57.4	282	1 PPNK_HALNI	P06079 lysobacter
14	39	57.4	523	1 UL21_PWN3	Q00703 salmonella
15	39	57.4	907	1 NUOG_SALTY	P33900 salmonell
16	38	55.9	266	1 GSJF_ERMCH	P24689 erwinia chr
17	38	55.9	268	1 CLCR_ERMCH	Q90788 gallus gall
18	38	55.9	294	1 NKZE_CHICK	P26490 astasia lon
19	38	55.9	432	1 PMS3_VIBCH	Q90788 gallus gall
20	38	55.9	432	1 RBL_ASTLO	P26490 astasia lon
21	38	55.9	626	1 Y124_DROME	Q24114 drosophila
22	38	55.9	733	1 Y124_DROME	Q24114 drosophila
23	38	55.9	1048	1 P100_HOMAN	Q9upw6 homo sapien
24	37.5	55.1	466	1 KRUP_DROME	P08318 human cytom
25	37	54.4	269	1 EL2_BOVIN	P07247 drosophila
26	37	54.4	436	1 RBL_EUGVI	Q29451 bos taurus
27	37	54.4	2441	1 CBP_MOUSE	Q45485 euglena vir
28	37	54.4	2672	1 GCN1_YEAST	P45485 euglena vir
29	36	52.9	167	1 OB_CANFA	P33892 mus musculu
30	36	52.9	214	1 RBL_MOUSE	O02720 canis fami
31	36	52.9	246	1 CB11_LYXSE	Q9cmz2 mus musculu
32	36	52.9	318	1 PTA_PARDE	P12360 lycopersico
33	36	52.9	397	1 POR_CHLRE	P39197 paracoccu
					Q39617 chlamydomon

34 36 52.9 429 1 YDIS_ECOLI
35 36 52.9 444 1 FXD4_MOUSE
36 36 52.9 451 1 ECFE_YERPE
37 36 52.9 746 1 CLCS_HUMAN
38 36 52.9 746 1 CLCS_MOUSE
39 36 52.9 746 1 CLCS_MOUSE
40 36 52.9 747 1 CLCS_RAT
41 36 52.9 747 1 CLC4_MOUSE
42 36 52.9 760 1 CLC3_CANPO
43 36 52.9 760 1 CLC3_MOUSE
44 36 52.9 760 1 CLC3_RAT
45 36 52.9 760 1 CLC4_HUMAN

F77337 escherichia
O60688 mus muscula
O82b59 versinia pe
P51795 homo sapien
Q9wv44 mus musculu
P51796 rattus norv
Q61418 mus musculu
P51794 rattus norv
Q91279 cavia porce
P51791 mus musculu
P51792 rattus norv
P51793 homo sapien

ALIGNMENTS

RESULT 1
CTRA_BOVIN
ID CTRA_BOVIN STANDARD; PRT; 245 AA.
AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]_TaxID=9913;
RX MEDLINE=69106266; PubMed=5764436;
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RA Brown J.R., Hartley B.S.;
RT "Location of disulfide bonds and active site of chymotrypsinogen A.";
RT The disulphide bridges of bovine chymotrypsinogen A.";
RN [2]
RP REVISION TO 102.
RA MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of chymotrypsin.";
RN [3]
RT Nature 221:337-340(1969).
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RN [4]
RN Nature 201:1284-1287(1964).
RP SEQUENCE, AND DISULFIDE BONDS
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Kluh I., Kostka V., Moravsek L., Prusik Z., Vanacek J.,
RA Keil B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RN [5]
RT Biochim. Biophys. Acta 130:543-546(1966).
RP ACTIVE SITE
RX MEDLINE=57181723; PubMed=5971785;
RA Smillie B., Hartley B.S.;
RT "Histidine residues in the active centres of some 'serine' proteinases"
RN [6]
RN Biochem. J. 101:232-241(1966).
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine proteinases. The structure of alpha-chymotrypsin.";
RN [7]
RN Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
RX MEDLINE=7017557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;

Wed Feb 12 11:59:41 2003

us-10-036-371-5.rap

"Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation.";
Biochemistry 9:1997-2009(1970).
[8]
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
MEDLINE=82078042; PubMed=6914398;
Cohen G.H., Silverton E.W., Davies D.R.;
Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
Comparison with other pancreatic serine proteases.";
J. Mol. Biol. 148:449-479(1981).
[9]
X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
MEDLINE=86011575; PubMed=4046030;
Tsukada H., Blow D.M.;
"Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
J. Mol. Biol. 184:703-711(1985).
RT
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- DATABASE: NAME=worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
CC
PR; A00952; KIBOA.
DR PDB; 2CGA; 15-APR-90.
DR PDB; 2CHA; 31-MAY-94.
DR PDB; 4CHA; 29-OCT-94.
DR PDB; 5CHA; 18-OCT-97.
DR PDB; 6CHA; 18-OCT-97.
DR PDB; 1CGI; 27-JAN-84.
DR PDB; 1CHJ; 16-JUL-88.
DR PDB; 2CGH; 31-MAY-84.
DR PDB; 3CGH; 15-OCT-92.
DR PDB; 4GCH; 15-OCT-90.
DR PDB; 5GCH; 15-OCT-90.
DR PDB; 6GCH; 15-OCT-90.
DR PDB; 7GCH; 15-OCT-90.
DR PDB; 8GCH; 15-JUL-93.
DR PDB; 1GCT; 15-OCT-91.
DR PDB; 2GCT; 15-OCT-91.
DR PDB; 3GCT; 15-OCT-91.
DR PDB; 1AGC; 31-OCT-93.
DR PDB; 1GMC; 31-OCT-93.
DR PDB; 1GMD; 31-OCT-93.
DR PDB; 1CGI; 30-APR-94.
DR PDB; 1CGJ; 30-APR-94.
DR PDB; 1GCD; 22-JUN-94.
DR PDB; 1GHA; 22-JUN-94.
DR PDB; 1GHB; 22-JUN-94.
DR PDB; 1GHH; 30-SEP-94.
DR PDB; 2GMT; 01-NOV-94.
DR PDB; 1AB9; 20-APR-97.
DR PDB; 1AFQ; 17-SEP-97.
DR PDB; 1CAH; 23-JUL-97.
DR PDB; 1CHW; 23-JUL-97.
DR PDB; 1VGC; 12-NOV-97.
DR PDB; 2VGC; 12-NOV-97.
DR PDB; 4VGC; 12-NOV-97.
DR PDB; 1HJN; 14-JAN-98.
DR MEROPS; S01.001.-.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
3D-structure.
KW CHYMOTRYPSIN A, A CHAIN.
FT CHAIN 1 13
CHAIN 16 146 CHYMOTRYPSIN A, B CHAIN.

CHYMOTRYPSIN A, C CHAIN.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.

149 245
CHAIN 57
ACT_SITE 102
ACT_SITE 195
ACT_SITE 122
DISULFID 1
DISULFID 42
DISULFID 58
DISULFID 136
DISULFID 168
DISULFID 191
DISULFID 220
HELIX 12
HELIX 20
HELIX 21
HELIX 24
HELIX 25
TURN 28
TURN 29
TURN 30
TURN 34
STRAND 40
STRAND 46
STRAND 51
STRAND 54
TURN 57
TURN 58
TURN 62
TURN 63
STRAND 65
STRAND 68
TURN 69
TURN 72
TURN 74
TURN 81
TURN 90
TURN 92
TURN 93
TURN 95
TURN 96
TURN 100
TURN 101
TURN 104
TURN 108
STRAND 122
STRAND 127
TURN 132
TURN 133
TURN 135
TURN 140
TURN 146
TURN 148
STRAND 154
STRAND 154
STRAND 156
STRAND 163
HELIX 165
HELIX 168
TURN 169
HELIX 172
HELIX 173
STRAND 175
STRAND 180
STRAND 184
TURN 192
TURN 193
TURN 195
TURN 196
STRAND 198
TURN 204
TURN 205
STRAND 206
TURN 216
TURN 218
TURN 219
TURN 222
TURN 223
STRAND 225
STRAND 230
HELIX 231
HELIX 233
TURN 234
TURN 234
HELIX 235
SEQUENCE 245
AA; 25666 MW; 91A9P28E2F3E3142 CRC64;

Query Match 100.0%; Score 68; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVPAIQPVLSGL 13
DB 1 GQVPAIQPVLSGL 13

RESULT 2
ID_CTRB_BOVIN STANDARD; PRT; 245 AA.
AC P00767;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DB Chymotrypsinogen B (EC 3.4.21.1).
OS Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RD
RX MECHANISM: DISULFIDE BONDS, AND ACTIVE SITE.
RY SWISS-PROT: P00766; PubMed=5649671;
RA Sailland, R. 1993. The structure of the chymotrypsinogen B compared with chymotrypsinogen A and
RT trypsinogen. J. Biol. Chem. 268:343-346(1993).
RL Nature 218:343-346(1996).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
CC Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHW.html".
DR PIR; A00953; KYBOB.
DR HSSP; P00766; IACB.
DR MEROPS; S01152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00069; Serprotease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00201; Tryp_Spc_1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
KW CHAIN 1 13
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 146 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 149 245 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 102 102 CHARGE RELAY SYSTEM.
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM.
FT DISULFID 42 122 CHARGE RELAY SYSTEM.
FT DISULFID 136 208
FT DISULFID 168 192
FT DISULFID 191 220
SQ SEQUENCE 245 AA; 25755 WW; 678016446F5FEB5 CRC64;
Query Match 100.0%; Score 68; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGVPAIQPVLSGL 13
DB 1 CGVPAIQPVLSGL 13
RESULT 3
ID_CTR2_CANFA STANDARD; PRT; 263 AA.
AC P04813;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RD
RX SEQUENCE FROM N.A.
RY MEDLINE=8417023; PubMed=6594866;
RA Pinsky S.D., LaGeorge K.S., Luc V., Scheele G.;
RT "Identification of canine pancreatic precchymotrypsinogen 2
RT sequence determination of canine pancreatic precchymotrypsinogen 2
RT mRNA."
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
CC Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use European Bioinformatics Institute. There are no restrictions on its
CC modified and published statements as long as its content is in no way
CC entitles requires a statement of agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01173; AAAJ0841.1;
DR PIR; A21195; A21195.
DR HSSP; P00766; IACB.
DR MEROPS; S01152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00069; Serprotease_Try.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00201; Tryp_Spc_1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
KW CHAIN 1 18
FT CHAIN 1 18 CHYMOTRYPSINOGEN 2.
FT CHAIN 19 263 CHYMOTRYPSIN 2, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN 2, B CHAIN.
FT ACT_SITE 167 263 CHYMOTRYPSIN 2, C CHAIN.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM.
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM.
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 184 219 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
FT DISULFID 263 263 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27787 MW; 2AF449D813B961 CRC64;
Query Match 100.0%; Score 68; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGVPAIQPVLSGL 13
DB 19 CGVPAIQPVLSGL 31
RESULT 4
ID_CTRB_HUMAN STANDARD; PRT; 263 AA.
AC P1753B;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RD
RX SEQUENCE FROM N.A.
RY MEDLINE=8417023; PubMed=2917002;
RA Tomita N., Matsubara K., Horii A., Doi S., Yokouchi H., Ogawa M.,
RY Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
RT precchymotrypsinogen cDNA."
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
CC Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

```

Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

```

CC CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa.
CC CC Phe-|-Xaa, Leu-|-Xaa.
CC CC -!- SUBCELLULAR LOCATION: Extracellular.
CC CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: M24400; AA52128.1; -.
CC CC EMBL: BC005385; AA905385.1; -.
CC CC DR PIR: A31299; A31299.
CC CC DR HSP: P00766; 1CHG.
CC CC DR MEROPS: S01.152; -.
CC CC DR GENE: HGNC:3521; CTRB1.
CC CC DR MIM: 118890; -.
CC CC DR InterPro: IPR001314; Chymotrypsin.
CC CC DR InterPro: IPR001254; Ser_protease_Try.
CC CC DR Pfam: PF00089; TrypSPC.
CC CC DR PRINTS: PR00722; CHYMOTRYPSIN.
CC CC DR SMART: SM00020; TRYPSIN_DOM; 1.
CC CC DR PROSITE: PS00240; TRYPSIN_HIS; 1.
CC CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
CC CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
CC CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
CC CC DR HydroLase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
CC CC SIGNAL 1 18 CHYMOTRYPSINOGEN B.
CC CC FT CHAIN 19 263 CHYMOTRYPSIN B, A CHAIN.
CC CC FT CHAIN 19 31 CHYMOTRYPSIN B, B CHAIN.
CC CC FT CHAIN 34 164 CHYMOTRYPSIN B, C CHAIN.
CC CC FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
CC CC FT ACT SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT ACT SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT ACT SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT DISULFID 19 140 BY SIMILARITY.
CC CC FT DISULFID 60 76 BY SIMILARITY.
CC CC FT DISULFID 154 219 BY SIMILARITY.
CC CC FT DISULFID 186 200 BY SIMILARITY.
CC CC FT DISULFID 209 238 BY SIMILARITY.
CC CC SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A4908B701 CRC64;

Query Match 92.6%; Score 63; DB 1; Length 263;
Best Local Similarity 92.3%; Pred. No. 0.0014; 1; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLSGL 13
DB 19 CGVPAIHPLVTLGL 31

RESULT 5
CTRB_RAT STANDARD; PRT; 263 AA.
AC P07138;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2003 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
CN CTRB1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054881; PubMed=6209274;
RA Bell G.I.; Quinto C.; Quiroga M.; Valenzuela P.; Craik C.S.;
RA Rutter W.J.;
RA "Isolation and sequence of a rat chymotrypsin B gene.";
RL J. Biol. Chem. 259:14265-14270(1984).

```

```

CC CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa.
CC CC Phe-|-Xaa, Leu-|-Xaa.
CC CC -!- SUBCELLULAR LOCATION: Extracellular.
CC CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: K02298; AAA98732.1; -.
CC CC DR PIR: A2658; KYTB.
CC CC DR HSP: P00766; 1CHG.
CC CC DR MEROPS: S01.152; -.
CC CC DR InterPro: IPR001314; Chymotrypsin.
CC CC DR InterPro: IPR001254; Ser_protease_Try.
CC CC DR Pfam: PF00089; trypsin; 1.
CC CC DR PRINTS: PR00722; CHYMOTRYPSIN.
CC CC DR SMART: SM00020; TRYPSIN_DOM; 1.
CC CC DR PROSITE: PS00240; TRYPSIN_HIS; 1.
CC CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
CC CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
CC CC DR HydroLase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
CC CC SIGNAL 1 18 CHYMOTRYPSINOGEN B.
CC CC FT CHAIN 19 263 CHYMOTRYPSIN B, A CHAIN.
CC CC FT CHAIN 19 31 CHYMOTRYPSIN B, B CHAIN.
CC CC FT CHAIN 34 164 CHYMOTRYPSIN B, C CHAIN.
CC CC FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
CC CC FT ACT SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT ACT SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT ACT SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT DISULFID 19 140 BY SIMILARITY.
CC CC FT DISULFID 60 76 BY SIMILARITY.
CC CC FT DISULFID 154 219 BY SIMILARITY.
CC CC FT DISULFID 186 200 BY SIMILARITY.
CC CC FT DISULFID 209 238 BY SIMILARITY.
CC CC SQ SEQUENCE 263 AA; 27849 MW; ACAFBACF8C4DA6D CRC64;

Query Match 89.7%; Score 61; DB 1; Length 263;
Best Local Similarity 84.6%; Pred. No. 0.0031;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLSGL 13
DB 19 CGVPTIQPVLTGL 31

RESULT 6
CTRB_HUMAN STANDARD; PRT; 264 AA.
AC P40313;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin-like protease CTRB-1 precursor (EC 3.4.21.-).
CN CTRB1 OR CTRB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94093544; PubMed=8268911;
RA Larsen F.; Solheim J.; Kristensen T.; Kolasa A.B.; Prydz H.;
RA "A tight cluster of five unrelated human genes on chromosome
RA 1q42.1.";
RL Hum Mol Genet. 2:1589-1595(1993).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X71874; CA507010.1; -
 CC EMBL; X71877; CA507011.1; -
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.256; -
 CC Gensu; H0NC:2524; CTRL.
 CC MIM; 114888; -

CC InterPro; IP001314; Chymotrypsin.
 CC InterPro; IP001254; Ser_Protease_Try.
 CC PRINTS; P00661; trypsin; 1.
 CC PROSITE; PS00204; CHYMOTRYPSIN.
 CC PROSITE; PS0240; TRYPSIN; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
 CC SIGNAL 1 18 POTENTIAL.

FT PROPEP 19 33 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 19 33 CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
 FT ACT_SITE 34 264 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 19 141 BY SIMILARITY.
 FT DISULFID 60 76 BY SIMILARITY.
 FT DISULFID 135 220 BY SIMILARITY.
 FT DISULFID 187 201 BY SIMILARITY.
 FT DISULFID 210 233 BY SIMILARITY.
 FT SIGNAL 264 AA; 28002 MW; 3F2ZF02FA6D84 CRC64;

Query Match 72.1%; Score 49; DB 1; Length 264;
 Best Local Similarity 72.7%; Pred. No. 0.95;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLIS 11
 ||:|||||
 DB 19 CGIPAIPALIS 29

RESULT 7
 ID CTRE GADMO STANDARD; PRT; 263 AA.
 AC P47736;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsin A precursor (EC 3.4.21.1).
 OS Gadus morhua (Atlantic cod).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 CC NCBI_TaxID=8049;
 CC L1 ||:|||||
 CC SEQUENCE FROM N.A.
 CC TISSUE=Pyloric caeca;
 CC MEDLINE=3436860; PubMed=8086467;
 CC Guzman J.B., Oskarsson S., Eakin A.E., Craik C.S.,
 CC Bjarnason J.B., Bjarnason J.B.;
 CC "Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";
 CC Biochim. Biophys. Acta 1219:211-214 (1994).
 CC (2)

RP SEQUENCE OF 19-30 AND 34-49.
 RC TISSUE=Pyloric caeca;
 RC MEDLINE=92111252; PubMed=1764912;
 RA Angelsson B., Bjarnason J.B.;
 RA "Structural and kinetic properties of chymotrypsin from Atlantic cod
 RT (Gadus morhua). Comparison with bovine chymotrypsin";
 RT

CC Comp. Biochem. Physiol. 99B:327-335(1991).
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-[Xaa, Trp]-[Xaa,
 CC Glu]-Xaa.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X78490; CA55242.1; -
 CC HSP; P00661; ICHG.
 CC MEROPS; S01.256; -

CC InterPro; IP001314; Chymotrypsin
 CC InterPro; IP001254; Ser_Protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM0020; TRYD_SPC; 1.
 CC PROSITE; PS0240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 CC SIGNAL 1 18

FT CHAIN 19 263 CHYMOTRYPSIN A.
 FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 19 140 BY SIMILARITY.
 FT DISULFID 150 216 BY SIMILARITY.
 FT DISULFID 186 200 BY SIMILARITY.
 FT DISULFID 209 238 BY SIMILARITY.
 FT DISULFID 21 21 R-S-C (IN REF. 2).
 FT CONFLICT 25 25 S->Q (IN REF. 2).
 FT CONFLICT 29 29 T->S (IN REF. 2).
 FT CONFLICT 44 44 S->T (IN REF. 2).
 FT CONFLICT 46 46 S->Y (IN REF. 2).
 FT SIGNAL 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

Query Match 69.1%; Score 47; DB 1; Length 263;
 Best Local Similarity 66.7%; Pred. No. 0.78;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLIS 12
 ||:|||||
 DB 19 CGIPAIPALIS 30

RESULT 8
 ID CTRE GADMO STANDARD; PRT; 245 AA.
 AC P0646;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsin B (EC 3.4.21.1).
 OS Gadus morhua (Atlantic cod).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 CC NCBI_TaxID=8049;
 CC L1 ||:|||||

RP SEQUENCE.
 RC TISSUE=Pyloric caeca;
 RC MEDLINE=96419045; PubMed=8841380;
 RA Leth-Larsen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M.,
 RA Hoejrup P.;
 RA "Structure of chymotrypsin variant B from Atlantic cod, Gadus
 RT morhua.";

biochim. Biophys. Acta 1297:49-56(1996).

SEQUENCE OF 1-12 AND 16-31.

SEQUENCE-Polioviric caeca;

MEDLINE=92011252; PubMed=1764912;

Accession B. Bjarnason J.B.;

"Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua). Comparison with bovine chymotrypsin.";

Comp. Biochem. Physiol. 99B:327-335(1991).

Catalytic Activity: Preferential cleavage: Tyr-[Xaa, Trp]-Xaa, Phe]-Xaa, Leu]-Xaa.

SUBCELLULAR LOCATION: Extracellular.

SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

HSP: P00766; 1CHG.

MEROPS: S01.152; --.

InterPro: IPR001314; Chymotrypsin.

IncePro: IPR001254; Ser protease_Try.

Pfam: PF00089; trypsin_1.

PRINTS: PR00722; CHYMOTRYPSIN.

SMART: SM00020; TRY1_SPC: 1.

PROSITE: PS0240; TRYPSIN_DOM; 1.

PROSITE: PS00134; TRYPSIN_HIS; 1.

PROSITE: PS00135; TRYPSIN_SER; 1.

Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.

CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.

CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.

ACT SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).

DISULFID 1 121 BY SIMILARITY.

DISULFID 42 58 BY SIMILARITY.

DISULFID 135 201 BY SIMILARITY.

DISULFID 167 182 BY SIMILARITY.

DISULFID 191 220 BY SIMILARITY.

CONFLICT 9 11 QVT -> VIS (JUN REF. 2).

CONFLICT 26 29 S -> Y (JUN REF. 2).

CONFLICT 28 29 EW -> Y (JUN REF. 2).

SEQUENCE 245 AA; 26260 MW; 74FE0D045517AB02 CRC64;

Query Match 67.6%; Score 46; DB 1; Length 245;

Best Local Similarity 66.7%; Pred. No. 1,1;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQVPAIQPVLSG 12

DB 1 CQSPALQPVVTS 12

RESULT 9

REL_CHLRE CHLRE STANDARD; PRT; 475 AA.

ID HE_HEMPU STANDARD; PRT; 475 AA.

AC P91953; 1998 (Rel. 01, Created)

DT 21-JUL-1998 (Rel. 01, Last sequence update)

DE 20-MAY-2000 (Rel. 39, Last annotation update)

DE Ribulose biphosphate carboxylase large chain precursor (EC 4.1.1.39)

DE (RuBisCO large subunit).

GN RECL

GN Chlamydomonas reinhardtii.

OS Chloroplast.

OS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonadae.

OC NCBI_TaxID=3055;

OX [1]

OX SEQUENCE FROM N.A.

RP MEDLINE=43189072; PubMed=6302265;

RX Dron M., Rahite M., Rochaix J.-D.;

RT "Sequence of the chloroplast DNA region of Chlamydomonas reinhardtii containing the gene of the large subunit of ribulose biphosphate carboxylase and parts of its flanking genes.";

RL J. Mol. Biol. 162:775-793(1982).

RL [2]

biochim. Biophys. Acta 1297:49-56(1996).

SEQUENCE OF 1-12 AND 16-31.

SEQUENCE-Polioviric caeca;

MEDLINE=92011252; PubMed=1764912;

Accession B. Bjarnason J.B.;

"Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua). Comparison with bovine chymotrypsin.";

Comp. Biochem. Physiol. 99B:327-335(1991).

Catalytic Activity: Preferential cleavage: Tyr-[Xaa, Trp]-Xaa, Phe]-Xaa, Leu]-Xaa.

SUBCELLULAR LOCATION: Extracellular.

SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

HSP: P00766; 1CHG.

MEROPS: S01.152; --.

InterPro: IPR001314; Chymotrypsin.

IncePro: IPR001254; Ser protease_Try.

Pfam: PF00089; trypsin_1.

PRINTS: PR00722; CHYMOTRYPSIN.

SMART: SM00020; TRY1_SPC: 1.

PROSITE: PS0240; TRYPSIN_DOM; 1.

PROSITE: PS00134; TRYPSIN_HIS; 1.

PROSITE: PS00135; TRYPSIN_SER; 1.

Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.

CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.

CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.

ACT SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).

DISULFID 1 121 BY SIMILARITY.

DISULFID 42 58 BY SIMILARITY.

DISULFID 135 201 BY SIMILARITY.

DISULFID 167 182 BY SIMILARITY.

DISULFID 191 220 BY SIMILARITY.

CONFLICT 9 11 QVT -> VIS (JUN REF. 2).

CONFLICT 26 29 S -> Y (JUN REF. 2).

CONFLICT 28 29 EW -> Y (JUN REF. 2).

SEQUENCE 245 AA; 26260 MW; 74FE0D045517AB02 CRC64;

Query Match 67.6%; Score 46; DB 1; Length 245;

Best Local Similarity 66.7%; Pred. No. 1,1;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQVPAIQPVLSG 12

DB 1 CQSPALQPVVTS 12

RESULT 9

REL_CHLRE CHLRE STANDARD; PRT; 475 AA.

ID HE_HEMPU STANDARD; PRT; 475 AA.

AC P91953; 1998 (Rel. 01, Created)

DT 21-JUL-1998 (Rel. 01, Last sequence update)

DE 20-MAY-2000 (Rel. 39, Last annotation update)

DE Ribulose biphosphate carboxylase large chain precursor (EC 4.1.1.39)

DE (RuBisCO large subunit).

GN RECL

GN Chlamydomonas reinhardtii.

OS Chloroplast.

OS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonadae.

OC NCBI_TaxID=3055;

OX [1]

OX SEQUENCE FROM N.A.

RP MEDLINE=43189072; PubMed=6302265;

RX Dron M., Rahite M., Rochaix J.-D.;

RT "Sequence of the chloroplast DNA region of Chlamydomonas reinhardtii containing the gene of the large subunit of ribulose biphosphate carboxylase and parts of its flanking genes.";

RL J. Mol. Biol. 162:775-793(1982).

RL [2]

biochim. Biophys. Acta 1297:49-56(1996).

SEQUENCE OF 1-12 AND 16-31.

SEQUENCE-Polioviric caeca;

MEDLINE=92011252; PubMed=1764912;

Accession B. Bjarnason J.B.;

"Structural and kinetic properties of chymotrypsin from Atlantic cod (Gadus morhua). Comparison with bovine chymotrypsin.";

Comp. Biochem. Physiol. 99B:327-335(1991).

Catalytic Activity: Preferential cleavage: Tyr-[Xaa, Trp]-Xaa, Phe]-Xaa, Leu]-Xaa.

SUBCELLULAR LOCATION: Extracellular.

SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

HSP: P00766; 1CHG.

MEROPS: S01.152; --.

InterPro: IPR001314; Chymotrypsin.

IncePro: IPR001254; Ser protease_Try.

Pfam: PF00089; trypsin_1.

PRINTS: PR00722; CHYMOTRYPSIN.

SMART: SM00020; TRY1_SPC: 1.

PROSITE: PS0240; TRYPSIN_DOM; 1.

PROSITE: PS00134; TRYPSIN_HIS; 1.

PROSITE: PS00135; TRYPSIN_SER; 1.

Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.

CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.

CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.

ACT SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).

DISULFID 1 121 BY SIMILARITY.

DISULFID 42 58 BY SIMILARITY.

DISULFID 135 201 BY SIMILARITY.

DISULFID 167 182 BY SIMILARITY.

DISULFID 191 220 BY SIMILARITY.

CONFLICT 9 11 QVT -> VIS (JUN REF. 2).

CONFLICT 26 29 S -> Y (JUN REF. 2).

CONFLICT 28 29 EW -> Y (JUN REF. 2).

SEQUENCE 245 AA; 26260 MW; 74FE0D045517AB02 CRC64;

Query Match 67.6%; Score 46; DB 1; Length 245;

Best Local Similarity 66.7%; Pred. No. 1,1;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQVPAIQPVLSG 12

DB 1 CQSPALQPVVTS 12

RESULT 9

REL_CHLRE CHLRE STANDARD; PRT; 475 AA.

ID HE_HEMPU STANDARD; PRT; 475 AA.

AC P91953; 1998 (Rel. 01, Created)

DT 21-JUL-1998 (Rel. 01, Last sequence update)

DE 20-MAY-2000 (Rel. 39, Last annotation update)

DE Ribulose biphosphate carboxylase large chain precursor (EC 4.1.1.39)

DE (RuBisCO large subunit).

GN RECL

GN Chlamydomonas reinhardtii.

OS Chloroplast.

OS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonadae.

OC NCBI_TaxID=3055;

OX [1]

OX SEQUENCE FROM N.A.

RP MEDLINE=43189072; PubMed=6302265;

RX Dron M., Rahite M., Rochaix J.-D.;

RT "Sequence of the chloroplast DNA region of Chlamydomonas reinhardtii containing the gene of the large subunit of ribulose biphosphate carboxylase and parts of its flanking genes.";

RL J. Mol. Biol. 162:775-793(1982).

RL [2]

biochim. Biophys. Acta 1297:49-56(1996).

RN [2]
 RP CLEAVAGE SPECIFICITY.
 RA MEDLINE=91283448; PubMed=1711895;
 RA Nomura K., Tanaka H., Kikkawa Y., Yamaguchi M., Suzuki N.:
 RA "Specificity of sea urchin hatching enzyme (envelysin) places it
 RT in the mammalian matrix metalloproteinase family.";
 RL Biochemistry 30:6115-6123(1991).
 RN [3]
 RP STEREO-SPECIFICITY.
 RA MEDLINE=93223852; PubMed=8467915;
 RA Nomura K., Suzuki N.:
 RA "Stereo-specific inhibition of sea urchin envelysin (hatching enzyme)
 RT by a synthetic autolysin inhibitor peptide with a cysteine-switch consensus
 RT sequence.";
 RL FEBS Lett. 321:84-88(1993).
 CC -1- FUNCTION: ALLOWS THE SEA URCHIN TO DIGEST THE PROTECTIVE ENVELOPE
 CC DERIVED FROM THE EGG EXTRACELLULAR MATRIX THUS ALLOWING THE SEA
 CC URCHIN TO SWIM FREELY.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: on the amino side of
 CC hydrophobic residues, -Leu-, -Ile-, -Phe, as well as -Tyr-.
 CC -1- SUBSTRATE: DURING HATCHING, THE 50 kDa MATURE ENZYME IS
 CC AUTOLYTICALLY CLEAVED TO PRODUCE A MAJOR 38 kDa AND A MINOR 15 kDa
 CC FORM WHICH MAY BE DISULFIDE LINKED. SUBSEQUENT CLEAVAGE OF THE 38
 CC kDa SPECIES YIELDS A MAJOR 38 kDa AND A MINOR 15 kDa
 CC -1- DEVELOPMENTAL STAGE: EMBRYO, BLASTULA STAGE. HIGHEST ACTIVITY AT
 CC 12.5 HRS EMBRYO STAGE.
 CC -1- DOMAIN: THERE ARE TWO DISTINCT DOMAINS IN THIS PROTEIN; THE
 CC CATALYTIC N-TERMINAL, AND THE C-TERMINAL WHICH IS INVOLVED IN
 CC SUBSTRATE SPECIFICITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A.
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR HMBL; AB000719; BAA19171.1; -.
 DR HSP; P03956; IGLG.
 DR MEROPS; M10.010; -.
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR001818; Matrxin.
 DR InterPro; IPR000130; Zn_Mtpeptidase.
 DR Pfam; PF00045; hemopexin; 4.
 DR Pfam; PF00413; Peptidase M10; 1.
 DR PRINTS; PR00138; MATRXIN.
 DR SMART; SM00120; HK; 4.
 DR SMART; SM00235; ZHM; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
 KW Autocatalytic cleavage; Signal; Calcium.
 FT SIGNAL 1 18
 FT PROPEP 19 169
 FT CHAIN 170 591
 FT CHAIN 170 591
 FT CHAIN 170 503
 FT CHAIN 504 591
 FT DOMAIN 62 76
 FT DOMAIN 336 380
 FT DOMAIN 380 589
 FT DOMAIN 448 451
 FT SITE 162 162
 FT MET 486 286
 FT ACT SITE 287 287
 FT METAL 490 490
 FT SITE 450 451
 FT SITE 503 503
 FT DISULFID 383 586
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L13200; AAA28194.1; -.
 DR PIR; S44899; S44899.
 DR WormPep; ZX1236.4; CE00531.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00078; fvc; 1.
 DR KEGG; K05000; Hatching enzyme.
 KW Hemopexin-like protein.
 FT SIGNAL 364 44; 42163 MW; B7BACEF8FE8DF80B CRC64;
 SQ SEQUENCE 364 44; 42163 MW; B7BACEF8FE8DF80B CRC64;

 FT CARBOHYD 129 129
 FT CARBOHYD 144 144
 FT CARBOHYD 578 578
 FT CARBOHYD 588 588
 FT CONFLICT 195 195
 FT CONFLICT 197 197
 FT CONFLICT 203 203
 FT CONFLICT 203 203
 FT CONFLICT 215 216
 FT CONFLICT 225 225
 FT CONFLICT 509 509
 FT CONFLICT 511 512
 FT CONFLICT 518 518
 FT CONFLICT 522 522
 FT CONFLICT 522 522
 SQ SEQUENCE 591 AA; 66126 MW; 5DCB448C6758C70D CRC64;

 Query Match 58.8%; Score 40; DB 1; Length 591;
 Best Local Similarity 58.3%; Pred. No. 27;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CUVPAQPVUSG 12
 DB 162 CUGVDLPFVVG 173

 RESULT 11
 Y084 CAEL
 ID Y084 CAEL STANDARD; PRF; 364 AA.
 AC P34620;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 42.2 kDa protein ZX1236.4 in chromosome III.
 GN ZX1236.4.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderae; Caenorhabditis.
 CC L3; TaxID=6239;
 RN L3; TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ba1501;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Anasacough R., Anderson K., Baynes C., Berke M.,
 RA Bonfield J., Burton J., Connors M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hacking I., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kerehav J., Kitzman J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimer M., Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Ropra A., Saunders D., Shawken R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden K.,
 RA Sultston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohlman P.;
 RL 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL Eegans 462 32-38(1994)
 CC -----
 CC -1- SIMILARITY: TO TRANSPOSON TI-2 OF AFRICAN MALARIA MOSQUITO.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L13200; AAA28194.1; -.
 DR PIR; S44899; S44899.
 DR WormPep; ZX1236.4; CE00531.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00078; fvc; 1.
 DR KEGG; K05000; Hatching enzyme.
 KW Hemopexin-like protein.
 FT SIGNAL 364 44; 42163 MW; B7BACEF8FE8DF80B CRC64;
 SQ SEQUENCE 364 44; 42163 MW; B7BACEF8FE8DF80B CRC64;

Wed Feb 12 11:59:41 2003

us-10-036-371-5.rsp

```

Query Match      58.1%; Score 39.5; DB 1; Length 364;
Seq. Local Similarity 56.2%; Pred. No. 20;
Matches          1; Mismatches 3; Indels 3; Gaps 1;

QY 1 CGVP--AIQPVLSGL 13
    |||||
DB 79 CGVPQGSVISPVLFQI 94

RESULT 12
ID PRLA_LYSEN STANDARD; PRT; 397 AA.
AC P00778;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-lytic protease precursor (EC 3.4.21.12) (Alpha-lytic
DE endopeptidase).
DE ALPHA-LP.
OS Lysobacter enzymogenes.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Lysobacter.
OX NCBI_TaxID=69;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29487;
RX MEDLINE=8934140; PubMed=3053694;
RA Epstein D.M.;
RT The alpha-lytic protease gene of Lysobacter enzymogenes. The
RT nucleotide sequence predicts a large prepro-peptide with homology to
RT pro-peptides of other chymotrypsin-like enzymes.";
RI 7; Biol. Chem. 263:16586-16590(1988).
RN [2]
RP REVISIONS.
RA Epstein D.M.;
RL Submitted (MAY-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29487;
RX MEDLINE=89170268; PubMed=3234766;
RA Silen J.L., McGrath C.N., Smith K.R., Agard D.A.;
RT Molecular analysis of the gene encoding alpha-lytic protease:
RT evidence for a preproenzyme.";
RL Gene 69:237-244(1988).
RN [4]
RP SEQUENCE OF 199-396.
RC STRAIN=ATCC 29487;
RX MEDLINE=71039222; PubMed=5482494;
RA Olson M.O.J., Nagabushan N., Dwiniel M., Smillie L.B.,
RT "Primary structure of alpha-lytic protease: a bacterial homologue of
RT the pancreatic serine proteases.";
RL Nature 226:438-442(1970).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=86074995; PubMed=117110;
RA Mayes G.D., Dalbaere L.T.J., James M.N.G.;
RT Molecular structure of the alpha-lytic protease from Myxobacter 495
RT at 2.8-A resolution.";
RL J. Mol. Biol. 131:743-775(1979).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=86011557; PubMed=3900416;
RA Fujinaga M., Dalbaere L.T.J., Brayer G.D., James M.N.G.;
RT "Refined structure of alpha-lytic protease at 1.7-A resolution.
RT Analysis of hydrogen bonding and solvent structure.";
RL J. Mol. Biol. 184:479-502(1985).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RC STRAIN=ATCC 29487;
RX MEDLINE=9893457; PubMed=9724517;
RA Peters R.J., Shiau A.K., Soni J.B., Anderson D.E., Tang G.,
RA Silen J.L., Agard D.A.;

```

```

"pro region C-terminus;protease active site interactions are critical
in catalyzing the folding of alpha-lytic protease.";
RL Biochemistry 37:12058-12067(1998).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=99023190; PubMed=9808037;
RA Sauter N.K., Mau T., Rader S.D., Agard D.A.;
RT "Structure of alpha-lytic protease complexed with its pro region.";
RL Nat. Struct. Biol. 5:945-950(1998).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins, especially bonds
CC adjacent to L-alanine and L-valine residues in bacterial cell
CC walls, elastin and other proteins.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A.
CC This SWISS-PROT entry is copyrighted by the EMBL, the European
CC Bioinformatics Institute, the Swiss Institute of Bioinformatics and the EMBL outstation -
CC between the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL: J04052; AAA25409.1; -;
DR EMBL: N22763; AAA74111.1; -;
DR PIR: A11721; TRYXBA
DR PDB: 1G8P; 15-OCT-89
DR PDB: 1G8A; 29-JAN-96
DR PDB: 1G8B; 29-JAN-96
DR PDB: 1G8C; 29-JAN-96
DR PDB: 1G8D; 29-JAN-96
DR PDB: 1G8E; 29-JAN-96
DR PDB: 1G8F; 29-JAN-96
DR PDB: 1G8G; 29-JAN-96
DR PDB: 1G8H; 29-JAN-96
DR PDB: 1G8I; 29-JAN-96
DR PDB: 1G8J; 29-JAN-96
DR PDB: 1G8K; 29-JAN-96
DR PDB: 1G8L; 29-JAN-96
DR PDB: 1G8M; 29-JAN-96
DR PDB: 2LPR; 15-JAN-93
DR PDB: 3LPR; 15-JAN-93
DR PDB: 5LPR; 15-JAN-93
DR PDB: 6LPR; 15-JAN-93
DR PDB: 7LPR; 15-JAN-93
DR PDB: 8LPR; 15-JAN-93
DR PDB: 9LPR; 15-JAN-93
DR PDB: 1P01; 15-APR-90
DR PDB: 1P02; 15-APR-90
DR PDB: 1P03; 15-APR-90
DR PDB: 1P04; 15-APR-90
DR PDB: 1P05; 15-APR-90
DR PDB: 1P06; 15-APR-90
DR PDB: 1P07; 15-APR-90
DR PDB: 1P11; 15-JAN-93
DR PDB: 1P12; 15-JAN-93
DR PDB: 1TAL; 01-APR-97
DR PDB: 2ULL; 07-JUL-97
DR PDB: 1B0Q; 12-AUG-98
DR PDB: 1Q04; 03-MAY-00
DR PDB: 1Q0W; 03-MAY-00
DR PDB: 1Q0X; 18-JUN-99
DR PDB: 2PRO; 27-APR-99
DR PDB: 3PRO; 27-APR-99
DR PDB: 4PRO; 18-MAY-99
DR MEROPS; S01.268; -;
DR InterPro; IPR004236; AL protease.
DR InterPro; IPR001316; Endopeptidase2A.
DR InterPro; IPR001254; Ser_protease_try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00861; ALYTCPTASE.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.

```

```

KW Hydrolase; Serine protease; 3D-structure; Zymogen; Signal.
FT SIGNAL 1 24
FT CHAIN 25 199
FT CH1 200 397
FT CH2 200 397
FT DISULFID 106 336
FT DISULFID 336 346
FT ACT SITE 235 235
FT ACT SITE 262 262
FT ACT SITE 342 342
FT ACT SITE 358 358
FT CONFLICT 171 171
FT STRAND 201 204
FT TURN 205 205
FT STRAND 207 210
FT TURN 211 213
FT STRAND 214 217
FT TURN 220 224
FT STRAND 222 225
FT STRAND 232 232
FT HELIX 232 232
FT TURN 239 240
FT STRAND 242 245
FT TURN 246 247
FT STRAND 248 257
FT STRAND 260 260
FT STRAND 263 268
FT TURN 270 271
FT STRAND 273 280
FT TURN 281 282
FT STRAND 283 286
FT STRAND 289 289
FT STRAND 295 296
FT STRAND 308 308
FT TURN 307 306
FT STRAND 307 324
FT TURN 323 324
FT STRAND 325 333
FT STRAND 338 338
FT TURN 339 340
FT TURN 342 343
FT STRAND 345 347
FT TURN 349 350
FT STRAND 352 350
FT TURN 364 365
FT STRAND 367 367
FT TURN 369 370
FT STRAND 373 375
FT HELIX 375 382
FT STRAND 384 384
FT TURN 391 391
FT STRAND 393 394
SQ SEQUENCE 397 AA; 41077 MW; 267FE6BEF57F33CB CRC64;

Query Match 58.1%; Score 39.5; DB 1; Length 397;
Best Local Similarity 40.0%; Pred. No. 22;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

QY 1 CGVPA-----IQPVLS 11
|||:|
|:|:|
Db 369 CGIPASORSSUFERLQPILS 388
|||:|
|:|:|

RESULT 13
ID PPNK HALN1 STANDARD; PRT; 282 AA.
AC Q9HNK7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23)
DE (Poly(P)/ATP NAD Kinase).
GN PPNK OR VNG190DC.

OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
NCBI_TaxID=64091;
RN 1
RX SEQUENCE FROM N.A.
MEDLINE=2050448; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Melias G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Dali T.A., Weitz R.,
RA Swartzell S., Weir D., Hall J.C., Danson M.J., Hough D.W.,
RA Lettshauser B., Keller K., Cruz K.E., Krebs M.P.,
RA Maddocks D.G., Jablonski P.E., Peck R.F., Pohlshroder M.,
RA Isenbarger T.A., Hou S., Daniels C.J., Dennis P.P., Jung K.-H.,
RA Alan M., Freitas T., Liang P., Riley M., Hood L., DasSarma S.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
RA "Genome sequence of Halobacterium species NRC-1.";
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC phosphate as a source of phosphorus (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + NAD(+) = ADP + NADP(+).
CC -!- COFACTOR: Requires divalent metal ions for activity (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright . It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not to be
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005089; XAG20093.1; -
DR InterPro; IP026254; ATP_NADK.
DR Pfam; PF01513; NADK; NAD_Kinase_1.
DR Trnasefame; Kinase; NAD_NADP; Complete proteome.
SQ SEQUENCE 282 AA; 29120 MW; 5DA1AF25147D7500 CRC64;

Query Match 57.4%; Score 39; DB 1; Length 282;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
|||:|
|:|:|
Db 207 GQADALPLVTGL 219
|||:|
|:|:|

RESULT 14
ID PPNK3 STANDARD; PRT; 523 AA.
AC Q00703;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE Protein U121 homolog.
DE Protein U121 homolog.
OS Pseudorabies virus (strain NIA-3) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=10349;
RN 1
RX SEQUENCE FROM N.A.
MEDLINE=93039656; PubMed=1331512;
RA Berman J., Kuman T.T., Pol J.J., Wagenaar F.F., de Wind N.N.;
RA "The pseudorabies virus homolog of the herpes simplex virus U121
RA gene product is a capsid protein which is involved in capsid
RA maturation.";
RL J. Virol. 66:7096-7103(1992)
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 U121,
CC HSV-1 40, EBV-4 U121, AND VZV 38
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

RN	14]	SEQUENCE FROM N.A.
RP		SPECIES=S.typhi; STRAIN=CT18;
RC		SPECIES=S.typhi; STRAIN=CT18;
RX		MEDLINE=21334947; PubMed=1157608;
RA		Parthill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RD		Charuher C., Wungall K.L., Bentley S.D., Holden M.T.G., Sebathia M.,

Query Match	Score 39;	Length 907;
Best Local Similarity	57.4%	
Pred. No. DB	63.6%	
Mismatches	2;	
Conservative	2;	
Indels	0;	
Gaps	0;	

QY 2 GVP AIQPVLSG 12
| : | | | : |
886 GMFGIAPVLAG 896

Wed Feb 12 11:59:41 2003

us-10-036-371-5.rsp

Page 11

Search completed: February 12, 2003, 10:23:17
Job time : 3.74527 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:16:06 : Search time 9.19701 Seconds
(without alignments)
291.248 Million cell updates/sec

Title: US-10-036-371-5

Perfect score: 68

Sequence: 1 CGVPAQPVLSGL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp nhc.*
- 8: sp organelle.*
- 9: sp phase.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp virus.*
- 16: sp bacterioph.*
- 17: sp archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	65	95.6	263	11	Q9DC86		Q9dc86 mus musculus
2	65	95.6	263	11	Q9DX88		Q9dx88 mus musculus
3	65	95.6	263	11	Q9CR35		Q9cr35 mus musculus
4	51	93.4	18	4	Q9UCF4		Q9ucf4 homo sapien
5	51	93.4	25	6	Q9TFL4		Q9trf4 sue scrofa
6	48	70.2	264	13	Q9PSL1		Q9psl1 gadus morhu
7	48	70.6	264	13	Q9PSL1		Q9psl1 gadus morhu
8	48	70.6	264	11	Q9D964		Q9d964 mus musculus
9	48	70.6	264	11	Q9D964		Q9d964 mus musculus
10	48	70.6	264	11	Q9D928		Q9d928 mus musculus
11	48	70.6	264	11	Q9ER05		Q9er05 mus musculus
12	47	69.1	23	13	Q9PSR0		Q9psr0 mus musculus
13	47	69.1	260	13	Q9W703		Q9w703 gadus morhu
14	46	67.6	263	13	Q9PMQ6		Q9pmq6 mus musculus
15	46	67.6	376	8	Q9XR04		Q9xr04 chlamydomon
16	46	67.6	376	8	Q9XR03		Q9xr03 chlamydomon

17	46	67.6	376	8	Q9XR02		Q9xr02 chlamydomon
18	46	67.6	376	10	Q65774		Q65774 chlamydomon
19	46	67.6	376	10	Q65777		Q65777 chlamydomon
20	46	67.6	376	10	Q9F505		Q9f505 streptomyce
21	44.5	65.4	384	4	Q9F505		Q9f505 streptomyce
22	44	64.7	384	13	Q9F5P2		Q9f5p2 yersinia pe
23	44	64.7	205	4	Q96P44		Q96p44 yersinia pe
24	44	64.7	270	5	Q96B74		Q96b74 homo sapien
25	43	63.2	391	2	Q98172		Q98172 rickettsella
26	42	61.8	376	8	Q98298		Q98298 chlamydomon
27	41	60.3	342	16	Q9YFD6		Q9yfd6 chlamydomon
28	40	58.8	153	8	Q95H37		Q95h37 eudorina mi
29	40	58.8	172	10	Q6W2X4		Q6w2x4 oryza sativ
30	40	58.8	195	10	Q946P4		Q946p4 caulobacter
31	40	58.8	255	16	Q944N6		Q944n6 caulobacter
32	40	58.8	326	16	Q82MQ2		Q82mq2 salmonella
33	40	58.8	376	8	Q24619		Q24619 playdiorina
34	40	58.8	376	8	Q24628		Q24628 pandorina u
35	40	58.8	376	8	Q24636		Q24636 pandorina u
36	40	58.8	376	8	Q24652		Q24652 pandorina u
37	40	58.8	376	8	Q24655		Q24655 pandorina u
38	40	58.8	376	8	Q24675		Q24675 pandorina u
39	40	58.8	376	8	Q9TJ39		Q9tj39 chlorogonium
40	40	58.8	376	8	Q9TJ39		Q9tj39 chlorogonium
41	40	58.8	376	8	Q9G149		Q9g149 chlorogonium
42	40	58.8	376	8	Q78227		Q78227 yamagishiella
43	40	58.8	376	8	Q9TH11		Q9th11 yamagishiella
44	40	58.8	376	8	Q98297		Q98297 chlorogonium
45	40	58.8	376	8	Q98300		Q98300 chlorogonium

ALIGNMENTS

RESULT 1

Q9DC86 Q9DC86 PRELIMINARY; PRT; 263 AA.
AC Q9DC86
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200080D09Rik protein.
GN 2200080D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPLEEN;
RC MEDLINE=21085660; PubMed=1117851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda Y.,
RA Satawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Kado K., Ozaki T., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kado K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Lewis S., Watson J., Utterback P., Kiehl J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita T., Hume D.A.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldrelli P., Bessie,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein S., Hill D., Hofmann M., Kazzarelli J., Gariboldi M.,
RA Gustincich S., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberti N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyoko-Oka K., Wang K.H., Weitz T., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:485-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILy.

```

DR EMBL; AK003060; BAB22539.1; -.
DR HSSP; P00766; 1GCT.
DR MEROPS; S01.152; -.
DR MGD; MGI:1913723; 2200008D09Rik.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PS00020; TRYPSIN_DOM.
DR SMART; SM00020; TRYPSIN_DOM.
DR PROSITE; PS00240; TRYPSIN_DOM.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HYDROLASE; Serine protease.
DR SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;

Query Match          95.6%; Score 65; DB 11; Length 263;
Best Local Similarity 92.3%; Pred. No. 0.0025;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLSGL 13
DB 19 CGVPAIQPVLTGL 31

RESULT 2
OY9DX8      PRELIMINARY;      PRT; 263 AA.
AC Q9DX8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; PubMed=1121785;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa H., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fetschmann M., Gaasterland T., Giesel C., King B., Kochiwa H.,
RA Kuehl L., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirli L., Stubi P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; AK007566; BAB25112.1; -.
CC HSSP; P00766; 1GCT.
CC MEROPS; S01.152; -.
CC MGD; MGI:1913723; 2200008D09Rik.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PS00020; TRYPSIN_DOM.
CC SMART; SM00020; TRYPSIN_DOM.
CC PROSITE; PS00240; TRYPSIN_DOM.
CC PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC HYDROLASE; Serine protease.
CC SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

```

```

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Hydrolase; Serine protease.
DR SQ SEQUENCE 263 AA; 27898 MW; C063BF8F905A92F CRC64;

Query Match          95.6%; Score 65; DB 11; Length 263;
Best Local Similarity 92.3%; Pred. No. 0.0025;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLSGL 13
DB 19 CGVPAIQPVLTGL 31

RESULT 3
OY9CR35      PRELIMINARY;      PRT; 263 AA.
AC Q9CR35;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; PubMed=1121785;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa H., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fetschmann M., Gaasterland T., Giesel C., King B., Kochiwa H.,
RA Kuehl L., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirli L., Stubi P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; AK008927; BAB25971.1; -.
CC EMBL; AK003079; BAB25253.1; -.
CC EMBL; AK007765; BAB25241.1; -.
CC EMBL; AK007815; BAB25280.1; -.
CC EMBL; AK008729; BAB25861.1; -.
CC EMBL; AK008888; BAB25954.1; -.
CC HSSP; P00766; 1GCT.
CC MEROPS; S01.152; -.
CC MGD; MGI:1913723; 2200008D09Rik.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PS00222; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM.
CC PROSITE; PS00240; TRYPSIN_DOM.
CC PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC HYDROLASE; Serine protease.
CC SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

```



```

Query Match          95.6%; Score 65; DB 11; Length 263;
Best Local Similarity 92.3%; Pred. No. 0.0025;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSGL 13
|||||:|||||:|
Db 19 CGVPAIQPVLTGL 31

RESULT 4
Q9UCF4 PRELIMINARY; PRT; 18 AA.
AC Q9UCF4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 14, Last annotation update)
DE Chymotrypsin isoenzyme CHT1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=93285747; PubMed=8509158;
RA Kim Y.J., Zhou Z., Hurtado J., Wood D.L., Choi A.S., Pescovitz M.D.,
RA Warfel K.A., Vandagriff J., Davis J.K., Kwon B.S.;
RA "IDM patients' sera recognize a novel 30-kD pancreatic autoantigen
RT related to chymotrypsinogen.";
RL Immunol. Invest. 22:219-227(1993).
SQ SEQUENCE 18 AA; 1826 MW; 38105D1037CA6837 CRC64;

Query Match          79.4%; Score 54; DB 4; Length 18;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVPAPQPVLSGL 13
|||||:|||||
Db 2 GVPAPQPVLSGL 13

RESULT 5
Q9TRL4 PRELIMINARY; PRT; 25 AA.
AC Q9TRL4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Chymotrypsinogen homolog (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RX MEDLINE=93016107; PubMed=1400470;
RA Gort S.U., Hamilton J.W., Cohn D.V.;
RA "Regulated, but not constitutive, secretory proteins bind porcine
RT chymotrypsinogen.";
RL J. Biol. Chem. 267:21595-21600(1992).
DR HSSP; P00766; 1ACB.
SQ SEQUENCE 25 AA; 2439 MW; F69A8003F1ACB136 CRC64;

Query Match          77.9%; Score 53; DB 6; Length 25;
Best Local Similarity 91.7%; Pred. No. 0.03;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVPAPQPVLSGL 13
|||||:|||||
Db 2 GVPAPQPVLSGL 13

RESULT 6
Q9PRS1 PRELIMINARY; PRT; 24 AA.
AC Q9PRS1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Chymotrypsin isoenzyme CHT1 (Fragment).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RX MEDLINE=96043258; PubMed=7584866;
RA Raaij A.J., Flengserud K., Stetten K.;
RA Chymotrypsin isoenzymes in Atlantic cod; differences in kinetics and
RT substrate specificities.
RL Comp. Biochem. Physiol. 112B:393-398(1995).
DR HSSP; P00766; 1ACB.
FT NON_TER 1 1
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2400 MW; 0A416ACA7B67F68D CRC64;

Query Match          70.6%; Score 48; DB 13; Length 24;
Best Local Similarity 75.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSG 12
|||||:|||||
Db 1 CGVPAIQPVSTG 12

RESULT 7
Q9WVQ4 PRELIMINARY; PRT; 261 AA.
AC Q9WVQ4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RX SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Sivasubava A.S., Kurokawa T.;
RA Japanese flounder cDNA for chymotrypsinogen 1.";
RL J. Biol. Chem. 273:1890-1893(1998).
CC -!- SIMILARITY BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BA882365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
RW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match          70.6%; Score 48; DB 13; Length 261;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGVPAIQPVLSG 12

```


RA Yamaguchi N.;
 RT Molecular cloning of rat chymopapsin *;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC EMBL: AB020757; BAB20287.1; -;
 DR HSSP: P00766; 4CHA.
 DR MEROPS: S01.256; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYPSIN_DOM; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolase: Serine protease.
 KW Hydrolyase: Serine protease.
 SQ SEQUENCE 264 AA; 2816 MW; F9ED50210FD3500E CRC64;

Query Match 70.6%; Score 48; DB 11; Length 264;

Best Local Similarity 81.8%; Pred. No. 2.2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLS 11
 ||||| |||

DB 19 CGVPAITPALS 29

RESULT 11

ID Q9ER05 PRELIMINARY; PRT; 264 AA.
 AC Q9ER05;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DR Chymotrypsin A CTBA-1.
 GN CTBL OR CTBA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Mitsui S., Yamaguchi N.;
 RT "Molecular cloning of mouse chymopapsin *";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC EMBL: AB016228; BAB20275.1; -;
 DR EMBL: AF213636; AAL11034.1; -;
 DR HSSP: P00766; 4CHA.
 DR MEROPS: S01.256; -;
 DR MGD: MGI:88558; Ctr1.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolase: Serine protease.
 KW Hydrolyase: Serine protease.
 SQ SEQUENCE 264 AA; 2815 MW; 1D979709A07056C2 CRC64;

Query Match 70.6%; Score 48; DB 11; Length 264;
 Best Local Similarity 81.8%; Pred. No. 2.2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLS 11
 ||||| |||

DB 19 CGVPAITPALS 29

RESULT 12

ID Q9PR50 PRELIMINARY; PRT; 23 AA.
 AC Q9PR50;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Chymotrypsin isoenzyme CH12 (fragment).
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=96043258; PubMed=7584866;
 RA Raai A.J., Flengsrud R., Sletten K.;
 RT "Chymotrypsin isoenzymes in Atlantic cod; differences in kinetics and
 RT substrate specificity.";
 RL Comp. Biochem. Physiol. 112B:393-398 (1995).
 FT NON_TER 1 1
 FT NON_TER 23 23
 FT SEQUENCE 23 AA; 2356 MW; A5EC2A7B7EB3F5E CRC64;

Query Match 69.1%; Score 47; DB 13; Length 23;
 Best Local Similarity 86.7%; Pred. No. 0.3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CGVPAIQPVLS 12

DB 1 CGVPAISPVITG 12

RESULT 13

ID Q9W7Q3 PRELIMINARY; PRT; 260 AA.
 AC Q9W7Q3;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Chymotrypsinogen 2.
 OS Paralichthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidae; Paralichthyidae; Paralichthys.
 OX NCBI_TaxID=8255;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RA Suzuki T., Srivastava A.S., Kurokawa T.;
 RT "Japanese flounder mRNA for chymotrypsinogen 2 *";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC EMBL: AB029754; BAB82366.1; -;
 DR HSSP: P00766; 1CHG.
 DR MEROPS: S01.152; 1.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR PRINTS: PR00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase: Serine protease.
 SQ SEQUENCE 260 AA; 27793 MW; 9F581044E22F78C0 CRC64;

```

Query Match          69.1%; Score 47; DB 13; Length 260;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 12
Db 16 CGSPAIPQVITG 27

RESULT 14
ID QSPM06 PRELIMINARY; PRT; 263 AA.
AC QSPM06;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 26, Last annotation update)
DE Chymotrypsin B precursor (EC 3.4.21.1).
GN Chyb.
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neoceratrygii; Teleostei; Euteleostei; Neoteleostei;
OC Neanthemoptera; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
RN 11 _taxid=8049;
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CHCA;
EX MEDLINE=2046433 PubMed=1011764;
RA Spilliaert R. Gendron et al.;
RT "Molecular Cloning of the cDNA for Cod Chymotrypsinogen B.";
RL Microb. Comp. Genomics 5:41-50(2000).
CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AJ242521; CAB43766.1; -.
DR HSPS; P00766; 1CHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SSR; 1.
DR Hydrolase; Serine protease; Signal.
FT SIGNAL 1 16
FT CHAIN 17 263 CHYMOTRYPSIN B.
SQ SEQUENCE 263 AA; 28175 MW; EF61B18A34EE57C CRC64;

Query Match          67.6%; Score 46; DB 13; Length 263;
Best Local Similarity 66.7%; Pred. No. 5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 12
Db 17 CGSPAIPQVITG 28

RESULT 15
ID QSPM06 PRELIMINARY; PRT; 376 AA.
AC QSPM06;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 32, Last annotation update)
DE Ribulose biphosphatase. carboxylase large chain (EC 4.1.1.39) (RuBisCO
DE large subunit) (Fragment).
GN RBCl.
OS Chlamydomonas mutabilis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=47905;

```

```

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=UTEX 578.
RA Morita E., Abe T., Tsuruki M., Fujiwara S., Sato N., Hirata A.,
RA Sorokite K., Noraki H.
RT "Role of pyrenes in the CO2-concentrating mechanism: comparative
RT morphology, physiology and molecular phylogenetic analysis of closely
RT related strains of Chlamydomonas and Chloromonas (Volvocales).";
RL planta 208:365-372(1999).
CC 1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
CC RIBULOSE 1,5-BISPHOSPHATE. THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION. AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE (BY SIMILARITY).
CC 1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
CC PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC 1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
CC PHOSPHO-D-GLYCERATE.
CC 1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
DR EMBL; AB022224; BAA82054.1; -.
DR InterPro; IPR000685; RUBISCO_large.
DR Pfam; PF00016; RUBISCO_large; 1.
DR PROSITE; PS02788; RUBISCO_large_N; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; 1.
KW Carbon dioxide fixation; Chloroplast; lyase; Monooxygenase;
KW Oxidoreductase; Photorespiration; Photosynthesis.
FT NON_TER 376 376
FT TER 376 376
SQ SEQUENCE 376 AA; 41774 MW; 845E743C6FB86A0 CRC64;

Query Match          67.6%; Score 46; DB 8; Length 376;
Best Local Similarity 58.3%; Pred. No. 7.1;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSG 12
Db 359 CGUGFWEVVASG 370

Search completed: February 12, 2003, 10:27:25
Job time : 11.197 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:18:40 ; Search time 2.79403 Seconds
(Without alignments)
136.898 Million cell updates/sec

Title: US-10-036-371-5

Perfect score: 68

Sequence: 1 CGVPAQPVLSGL 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PTCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	61.8	326	3	US-08-988-876-1
2	39.5	58.1	198	1	US-08-278-091-16
3	39.5	58.1	198	1	US-08-483-859-16
4	39.5	58.1	198	1	US-08-472-173-16
5	39.5	58.1	198	2	US-08-487-167-16
6	39.5	58.1	198	2	US-08-482-816-16
7	39.5	58.1	198	2	US-08-296-149-16
8	39.5	58.1	198	2	US-08-801-499-16
9	39.5	58.1	198	2	US-08-615-271-16
10	39.5	58.1	198	3	US-09-074-460-16
11	39.5	58.1	198	3	US-09-074-659-16
12	39.5	58.1	198	3	US-09-106-468-16
13	39.5	58.1	198	4	US-09-106-466A-16
14	39.5	58.1	198	4	US-09-106-467-16
15	38	57.4	525	1	US-08-681-129-2
16	38	55.9	268	1	US-08-270-584A-2
17	38	55.9	268	2	US-08-765-192-2
18	38	55.9	268	3	US-09-199-793-2
19	38	55.9	1048	4	US-09-171-699-10
20	37	54.1	2441	1	US-08-194-468-2
21	37	54.1	2441	3	US-08-961-739-2
22	37	54.1	2441	3	US-08-514-247A-8
23	36	52.9	70	2	US-08-935-450-12
24	36	52.9	146	4	US-08-914-375C-37
25	36	52.9	739	3	US-08-510-646B-33
26	36	52.9	837	2	US-08-474-067-7
27	36	52.9	837	2	US-08-474-068A-7

ALIGNMENTS

RESULT 1
US-08-988-876-1
; Sequence 1, Application US/08988876
; Patent No. 6063596
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hue, Henry
; TITLE OF INVENTION: PROTEIN COUPLED RECEPTORS ASSOCIATED
; TITLE OF INVENTION: WITH IMMUNE RESPONSE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PARCSCO for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/988,876
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0441 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT01
; CLONE: 364702
US-08-988-876-1

Sequence 6, Appli
Sequence 163, App
Sequence 317, App
Sequence 3728, Ap
Sequence 5, Appli
Sequence 4, Appli
Sequence 3497, Ap
Sequence 11, Appli
Sequence 11, Appli
Patent No. 15196194
Sequence 2, Appli
Sequence 8, Appli
Sequence 18, Appli
Sequence 3, Appli
Sequence 127, App
Sequence 13, Appli

```
Query Match 61.8%; Score 42; DB 3; Length 326;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 0; Gaps 0;

Oy 2 GVPAPQPVLSGL 13
Db 42 GDPELPQVLAGL 53

RESULT 2
US-08-278-091-16
; Sequence 16, Application US/08278091
; Patent No. 5656136
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION/DOCKET NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-495 MIS.vg
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-483-859-16
; Query Match 58.1%; Score 39.5; DB 1; Length 198;
; Best Local Similarity 40.0%; Pred. No. 23;
; Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

Oy 1 CGVPA-----IQPVLS 11
Db 170 CGIPASQRSSLPRLQPILS 189

RESULT 4
US-08-472-173-16
; Sequence 16, Application US/08472173
; Patent No. 5665353
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```



```

Oy 1 CGVPA-----IQPVLS 11
Db 170 CGIPASQSSLPFLRLOPILS 189

RESULT 7
US-08-296-149-16
; Sequence 16, Application US/08296149
; Patent No. 591297
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,149
; FILING DATE: 21-JUL-1994
; PRIORITY INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-801-499-16
; Query Match 58.1%; Score 39.5; DB 2; Length 198;
; Best Local Similarity 40.0%; Pred. No. 23;
; Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

Oy 1 CGVPA-----IQPVLS 11
Db 170 CGIPASQSSLPFLRLOPILS 189

RESULT 9
US-08-615-271-16
; Sequence 16, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271

```



```

; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-613-271-16

```

```

Query Match 58.1%; Score 39.5; DB 2; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

```

```

Qy 1 CGVPA-----IQPVLS 11
Db 170 CGIPASQSRSLFERLQPILS 189

```

RESULT 10

```

US-09-074-660-16
; Sequence 16, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:

```

```

; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada

```

```

; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-660-16

```

```

Query Match 58.1%; Score 39.5; DB 3; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

```

```

Qy 1 CGVPA-----IQPVLS 11
Db 170 CGIPASQSRSLFERLQPILS 189

```

RESULT 11

```

US-09-074-659-16
; Sequence 16, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:

```

```

; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada

```

```

; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,659
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; REFERENCE/DOCKET NUMBER: 1038-730 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-659-16

```

```

Query Match 58.1%; Score 39.5; DB 3; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

```

```

Qy 1 CGVPA-----IQPVLS 11

```

```

DB      170 CGIPASQSSIFERLQPILS 189          :|||:|
; NUMBER OF SEQUENCES: 23                      ;
; CORRESPONDENCE ADDRESS:                     ;
; ADDRESSEE: Sim & McBurney                   ;
; STREET: 6th Floor, 330 University Avenue     ;
; CITY: Toronto                               ;
; STATE: Ontario                              ;
; COUNTRY: Canada                             ;
; ZIP: M5G 1R7                                ;
; COMPUTER READABLE FORM:                     ;
; MEDIUM TYPE: Floppy disk                    ;
; COMPUTER: IBM PC compatible                  ;
; OPERATING SYSTEM: PC-DOS/MS-DOS             ;
; SOFTWARE: Patent In Release #1.0, Version #1.30 ;
; CURRENT APPLICATION DATA:                   ;
; APPLICATION NUMBER: US/09/106,466A           ;
; FILING DATE:                                 ;
; CLASSIFICATION: 514514                       ;
; PRIOR APPLICATION DATA:                     ;
; APPLICATION NUMBER: US 08/615,271            ;
; FILING DATE: 20-JUN-1996                    ;
; CLASSIFICATION: 514514                       ;
; ATTORNEY/AGENT INFORMATION:                 ;
; NAME: Stewart, Michael I                     ;
; REGISTRATION NUMBER: 24,973                  ;
; REFERENCE/DOCKET NUMBER: 1038-826           ;
; TELECOMMUNICATION INFORMATION:               ;
; TELEPHONE: (416) 595-1155                    ;
; TELEFAX: (416) 595-1163                      ;
; INFORMATION FOR SEQ ID NO: 16:                ;
; SEQUENCE CHARACTERISTICS:                    ;
; LENGTH: 198 amino acids                      ;
; TYPE: amino acid                             ;
; STRANDEDNESS: single                         ;
; TOPOLOGY: linear                            ;
; US-09-106-466A-16                           ;

Query Match      58.1%; Score 39.5; DB 4; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9;

OY   1   CGVPA-----IQPVLS 11
DB    170 CGIPASQSSIFERLQPILS 189
; Sequence 16 Application US/09106467
; Patent No. 6143540
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,466A
; FILING DATE:
; CLASSIFICATION: 514514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-826
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-466A-16

Query Match      58.1%; Score 39.5; DB 3; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

OY   1   CGVPA-----IQPVLS 11
DB    170 CGIPASQSSIFERLQPILS 189
; Sequence 16 Application US/09106466A
; Patent No. 6144125
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,468
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-468-16

Query Match      58.1%; Score 39.5; DB 3; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

OY   1   CGVPA-----IQPVLS 11
DB    170 CGIPASQSSIFERLQPILS 189
; Sequence 16 Application US/09106466A
; Patent No. 6147057
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; PROTEASE ACTIVITY

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-106-467-16

Query Match      58.1%; Score 39.5; DB 4; Length 198;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 0; Indels 9; Gaps 1;

QY 1 GQVPA-----IOPVLS 11
   |||||
DB 170 GGPASQSSLFERLPILS 189

RESULT 15
US-08-681-129-2
; Sequence 2, Application US/08681129
; Patent No. 7,318,954
; GENERAL INFORMATION:
; APPLICANT: M. J. Letier, Thomas Cristoph
; TITLE OF INVENTION: Pseudorabies virus vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,129
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,446
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 424
; APPLICATION NUMBER: EP 92.203.079.6
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus

```

```

; STRAIN: Kaplan
; US-08-681-129-2
Query Match      57.4%; Score 39; DB 1; Length 525;
Best Local Similarity 54.5%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVPAlOPVLSG 12
   |:::|
DB 187 GPGVRPPLSG 197

Search completed: February 12, 2003, 10:30:10
Job time : 3.79403 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:22:36 ; Search time 1.9791 Seconds
(without alignments)
167.821 Million cell updates/s

Title: US-10-036-371-5

perfect score: 68

Sequence: 1 CGVPAIQPVLSGL 13

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs. 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing. Minimum Match 08

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

Database :	Published Applications_Aa.*
1:	/cn2/6/prodata/1/pubpaa/US08_NEW_PUB.pcp.*
2:	/cn2/6/prodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3:	/cn2/6/prodata/1/pubpaa/US06_NEW_PUB.pcp.*
4:	/cn2/6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
5:	/cn2/6/prodata/1/pubpaa/US07_NEW_PUB.pcp.*
6:	/cn2/6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
7:	/cn2/6/prodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
8:	/cn2/6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
9:	/cn2/6/prodata/1/pubpaa/US09_NEW_PUB.pcp.*
10:	/cn2/6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
11:	/cn2/6/prodata/1/pubpaa/US10_NEW_PUB.pcp.*
12:	/cn2/6/prodata/1/pubpaa/US10_PUBCOMB.pcp.*
13:	/cn2/6/prodata/1/pubpaa/US60_NEW_PUB.pcp.*
14:	/cn2/6/prodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	68	100.0	13	12	US-10-036	371-5	Sequence 5, Appl
2	63	92.6	16	10	US-09-268	217-643	Sequence 643, App
3	93	92.6	16	10	US-09-268	217-643	Sequence 643, App
4	49	72.1	192	10	US-09-929	529-529	Sequence 529, App
5	40	58.8	13	12	US-10-036	371-5	Sequence 3, Appl
6	39	57.4	78	10	US-09-764	860-944	Sequence 944, App
7	39	57.4	168	9	US-09-862	540-34	Sequence 34, Appl
8	38	55.9	565	9	US-09-999	248-4	Sequence 4, Appl
9	37	54.4	24	9	US-09-774	639-283	Sequence 283, App
10	37	54.4	24	9	US-09-774	639-283	Sequence 283, App
11	37	54.4	51	10	US-09-866	761-36817	Sequence 36817, App
12	37	54.4	52	10	US-09-864	761-46426	Sequence 46426, A
13	37	54.4	99	10	US-09-764	860-305	Sequence 305, App
14	37	54.4	125	9	US-10-114	893-313	Sequence 313, App
15	37	54.4	127	10	US-09-866	761-33801	Sequence 33801, A
16	37	54.4	1827	9	US-09-772	363-261	Sequence 261, App
17	37	54.4	92	10	US-09-864	761-46426	Sequence 46426, A
18	37	54.4	4019	10	US-09-864	133-425	Sequence 425, App
19	37	54.4	4019	10	US-09-738	373-329	Sequence 329, App

ALIGNMENTS

```

RESULT 1
US/10/036-371-5
; STATEMENT 5, APPLICATION US/10036371
; PATENT NO. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; TITLE OF INVENTION: COSMETIC USE
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371

```

Query Match	100.0%;	Score 68;	DB 12;	Length 13;
Best Local Similarity	100.0%;	Pred. No. 1.1e-05;		
Matches 13:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

ON 1 SEPTEMBER 1971

100

90

80

70

60

50

40

30

20

10

0

0 1 2 3 4 5 6 7 8 9 10 11 12

12

```

RESULT 2
US-09-925-297-643
; Sequence 643, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, P
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925
; CURRENT FILING DATE: 2001-08-10

```

```

; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 643
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-297-643

Query Match          92.6%; Score 63; DB 10; Length 146;
Best Local Similarity 92.3%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSGL 13
Db 23 CGVPAIQPVLSGL 35

RESULT 3
; Sequence 96, Application US/09888615
; Patent No. US2002064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-888-615-96

Query Match          92.6%; Score 63; DB 10; Length 263;
Best Local Similarity 92.3%; Pred. No. 0.0019;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLSGL 13

```

```

Db 19 CGVPAIQPVLSGL 31

RESULT 4
; US-09-925-297-529
; Sequence 529, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS AND ANTIBODIES
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 529
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-297-529

Query Match          72.1%; Score 49; DB 10; Length 192;
Best Local Similarity 72.7%; Pred. No. 0.31;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGVPAIQPVLS 11
Db 15 CGIPAIKPAUS 25

RESULT 5
; US-10-036-371-3
; Sequence 3, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: ELMARSSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE OF INVENTION: COSMETIC USE
; FILE REFERENCE: 81691/28460
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (3)
; OTHER INFORMATION: R or S
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (7)
; OTHER INFORMATION: S or Q
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (9)
; OTHER INFORMATION: V or Q
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (10)
; OTHER INFORMATION: I or V
; US-10-036-371-3

```

```

Query Match      58.8%; Score 40; DB 12; Length 13;
Best Local Similarity 58.3%; Pred. No. 0.54;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CQVPAIQPVLSG 12
Db 1 CQVPAIXPXTG 12

RESULT 6
US-09-764-869-944
; Sequence 944, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1: 44
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-869-944

Query Match      57.4%; Score 39; DB 10; Length 78;
Best Local Similarity 53.8%; Pred. No. 5.6;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CQVPAIQPVLSGL 13
Db 15 CGIPAGDPAASSL 27

RESULT 7
US-09-862-540-34
; Sequence 34, Application US/09862540
; Patent No. US20030023992A1
; GENERAL INFORMATION:
; APPLICANT: Vogell, Gabriel
; TITLE OF INVENTION: No. US20030023992A1e1 G Protein-Coupled Receptors
; FILE REFERENCE: 00252US1
; CURRENT APPLICATION NUMBER: US/09/862,540
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,138
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/206,139
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/208,976
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 34
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-540-34

Query Match      57.4%; Score 39; DB 9; Length 168;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PAIQPVLSGL 13
Db 90 PELQPVLAGL 99

RESULT 8
US-09-999-248-4

```

```

; Sequence 4, Application US/09999248
; Patent No. US20020176852A1
; GENERAL INFORMATION:
; APPLICANT: Latch, J. David
; TITLE OF INVENTION: Mitogenic Oxygenase Regulators
; FILE REFERENCE: 05501-0180 43150-266489
; CURRENT APPLICATION NUMBER: US/09/999,248
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/249,305
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/251,364
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/289,172
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/289,537
; PRIOR FILING DATE: 2001-05-07
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4: 26
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-999-248-4

Query Match      55.9%; Score 38; DB 9; Length 565;
Best Local Similarity 58.3%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CQVPAIQPVLSG 12
Db 540 CGSPALAKVLKG 551

RESULT 9
US-09-774-639-283
; Sequence 283, Application US/09774639
; Patent No. US20030003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 283
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-283

Query Match      54.4%; Score 37; DB 9; Length 24;
Best Local Similarity 63.6%; Pred. No. 3.4;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GVPVPAIQPVLSG 12
Db 10 GVPVHPAGSG 20

RESULT 10
US-09-774-639-282
; Sequence 282, Application US/09774639
; Patent No. US20030003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; CURRENT FILING DATE: 2001-07-09

```

```

, PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
, PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
, NUMBER OF SEQ ID NOS: 371
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 282
, LENGTH: 48
, TYPE: PRT
, ORGANISM: Homo sapiens
US-09-774-639-282

```

```

Query Match      54.4%; Score 37; DB 9; Length 48;
Best Local Similarity 63.6%; Pred. No. 7.2;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy  2 GVPALQPVLSG 12
    |||||: ||
Db  20 GVPVHPAGSG 30

```

```

RESULT 11
US-09-864-761-36817
, Sequence 36817, Application US/09864761
, Patent No. US20020048763A1
, GENERAL INFORMATION:
, APPLICANT: Penn, Sharon G.
, APPLICANT: Rank, David R.
, APPLICANT: Hanzel, David K.
, APPLICANT: Chen, Wensheng
, TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
, TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
, FILE REFERENCE: Aemlica-X-1
, CURRENT APPLICATION NUMBER: US/09/864,761
, CURRENT FILING DATE: 2001-05-23
, PRIOR APPLICATION NUMBER: US 60/180,312
, PRIOR FILING DATE: 2000-02-04
, PRIOR APPLICATION NUMBER: US 60/207,456
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: US 09/632,366
, PRIOR FILING DATE: 2000-08-03
, PRIOR APPLICATION NUMBER: GB 24263.6
, PRIOR FILING DATE: 2000-10-04
, PRIOR APPLICATION NUMBER: US 60/236,359
, PRIOR FILING DATE: 2000-09-27
, PRIOR APPLICATION NUMBER: PCT/US01/00666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00665
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00668
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00663
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00662
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00661
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Annumax Sequence Listing Engine vers. 1.1
, SEQ ID NO 36817
, LENGTH: 51

```

```

, TYPE: PRT
, ORGANISM: Homo sapiens
, FEATURE:
, OTHER INFORMATION: MAP TO AL049733.3
, OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
, OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
, OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.93
, OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92
, OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.91
, OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
, OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
, OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
, OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95
, OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
, OTHER INFORMATION: EST_HUMAN HIT: AW961297.1, EVALUATE 2.00e-03
US-09-864-761-36817

```

```

Query Match      54.4%; Score 37; DB 10; Length 51;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy  2 GVPALQPVLSG 13
    |||||: ||
Db  10 GIPPIPMFGL 21

```

```

RESULT 12
US-09-864-761-46426
, Sequence 46426, Application US/09864761
, Patent No. US20020048763A1
, GENERAL INFORMATION:
, APPLICANT: Penn, Sharon G.
, APPLICANT: Rank, David R.
, APPLICANT: Hanzel, David K.
, APPLICANT: Chen, Wensheng
, TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
, TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
, FILE REFERENCE: Aemlica-X-1
, CURRENT APPLICATION NUMBER: US/09/864,761
, CURRENT FILING DATE: 2001-05-23
, PRIOR APPLICATION NUMBER: US 60/180,312
, PRIOR FILING DATE: 2000-02-04
, PRIOR APPLICATION NUMBER: US 60/207,456
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: US 09/632,366
, PRIOR FILING DATE: 2000-08-03
, PRIOR APPLICATION NUMBER: GB 24263.6
, PRIOR FILING DATE: 2000-10-04
, PRIOR APPLICATION NUMBER: US 60/236,359
, PRIOR FILING DATE: 2000-09-27
, PRIOR APPLICATION NUMBER: PCT/US01/00666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00667
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00664
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00669
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00665
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00668
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00663
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00662
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00661
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408

```



```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46426
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; SEQUENCE INFORMATION: MAP TO AC010251.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EST_HUMAN HIT: BE663509.1, EVALUATE 2.00e-04
US 09-864-761-46426

```

```

Query Match          54.4%; Score 37; DB 10; Length 52;
Best Local Similarity 63.3%; Pred. No. 7.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 CQVPAIQPVLS 11
| | | | |
DB 42 CTPPALMPVIS 52

```

```

RESULT 13
US-09-764-860-305
; Sequence 305; Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Eeren et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 305
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (40)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (79)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-305

```

```

Query Match          54.4%; Score 37; DB 10; Length 99;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 CQVPAIQPVLS 13
| | | | |
DB 74 GVPLKXPLXGL 85

```

```

RESULT 14
US-10-114-893-313
; Sequence 313; Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.

```

```

; APPLICANT: LeVallie, Edward R.
; APPLICANT: Collins-Macie, Lisa A.
; APPLICANT: Eeren, Chantal
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; PRIOR FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-313

Query Match          54.4%; Score 37; DB 9; Length 125;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVPAIQPVLSGL 13
| : | : | : |
DB 52 GLPALQPLTSL 63

RESULT 15
US-09-864-761-33801
; Sequence 33801; Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Menseheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,697
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33801
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009237.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEAT04281, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: B04042811, EVALUE 2.00e-40
; OTHER INFORMATION: EST_HUMAN HIT: B04527411, EVALUE 2.00e-45
; OTHER INFORMATION: SWISSPROT HIT: Q95157, EVALUE 9.00e-47
US-09-864-761-33801

Query Match 54.4%; Score 37; DB 10; Length 147;
Best Local Similarity 70.0%; Pred No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PAIQPVLSGL 13
Db 25 PELQPVLPGL 34

Search completed: February 12, 2003, 10:31:06
Job time : 2.9791 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:03:55 : Search time 13.4328 Seconds
(Without alignments)
196.395 Million cell updates/sec

Title: US-10-036-371-6

Perfect score: 112

Sequence: 1 INVGEAVPGSWPQVSIQD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	112	100.0	20	22	AA131578		Amino terminal seq
2	112	100.0	245	21	AA199596		Bovine chymotrypsin
3	109	97.3	20	20	AA193344		Bovine chymotrypsin
4	109	97.3	20	22	AA193935		N-terminal of a bo
5	109	97.3	20	22	AA193935		N-terminal of a bo
6	109	97.3	146	22	AA194791		Human pancreatic c
7	109	97.3	231	22	AA198401		Human chymotrypsin
8	109	97.3	263	23	AA192738		Amino acid sequenc
9	107	95.5	264	21	AA111711		Mouse serine prote
10	97	86.6	192	21	AA154077		Human pancreatic c

11	97	86.6	264	21	AA111710	Human serine prote
12	93	83.0	31	17	AA02588	PX, porcine pancre
13	93	83.0	23	13	AA190061	Sequence of human
14	93	83.0	23	13	AA190061	Porcine pancreatic
15	90	80.4	29	16	AA192621	Serum calcium lowe
16	90	80.4	268	17	AA190582	Rat caldesmon cont
17	88	78.6	242	7	AA190060	Sequence of human
18	88	78.6	270	8	AA190759	Human pancreas ela
19	88	78.6	270	23	AA197691	Human pancreas c
20	86	76.8	252	7	AA190062	Sequence of human
21	86	76.8	269	7	AA190062	Sequence of human
22	86	76.8	269	7	AA190062	Human elastase II
23	86	76.8	269	21	AA190062	Human pancreatic c
24	86	76.8	273	21	AA190062	Human pancreatic c
25	86	76.8	273	21	AA190062	Novel human diagno
26	86	76.8	273	21	AA190062	Novel human diagno
27	86	76.8	273	21	AA190062	Novel human diagno
28	84	75.0	283	21	AA190062	Human prostate-ass
29	84	75.0	283	21	AA190062	Human prostate-ass
30	84	75.0	283	21	AA190062	Human prostate-ass
31	84	75.0	283	21	AA190062	Human prostate-ass
32	84	75.0	283	21	AA190062	Human prostate-ass
33	84	75.0	283	21	AA190062	Human prostate-ass
34	84	75.0	283	21	AA190062	Human prostate-ass
35	84	75.0	283	21	AA190062	Human prostate-ass
36	84	75.0	283	21	AA190062	Human prostate-ass
37	84	75.0	283	21	AA190062	Human prostate-ass
38	84	75.0	283	21	AA190062	Human prostate-ass
39	84	75.0	283	21	AA190062	Human prostate-ass
40	84	75.0	283	21	AA190062	Human prostate-ass
41	84	75.0	283	21	AA190062	Human prostate-ass
42	84	75.0	283	21	AA190062	Human prostate-ass
43	84	75.0	283	21	AA190062	Human prostate-ass
44	84	75.0	283	21	AA190062	Human prostate-ass
45	84	75.0	283	21	AA190062	Human prostate-ass

ALIGNMENTS

RESULT 1

ID AAB31578 standard; peptide; 20 AA.

XX AAB31578;

XX AAB31578;

DT 20-APR-2001 (first entry)

DE Amino terminal sequence of bovine trypsin B chain.

XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;

XX Arthritis; inflamed joint; bursitis; osteoarthritis; fibromyalgia;

XX Rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;

XX Acne; eczema; facial seborrheic dermatitis; tendinitis; rash; psoriasis;

XX Athlete's foot; fungal infection; ulcer; navel infection; wrinkle;

XX Scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;

XX Fungal infection; autoimmune disease.

XX Bos sp.

XX W0200078332-A2.

XX 28-DEC-2000.

XX 15-JUN-2000; 2000NO-ISO0005.

XX 18-JUN-1999; 99IS-0005086.

XX (BJAR/) BJARNASON J B.

XX Bjarnason JB;

XX

DR WPI; 2001-091493/10.

XX Fish serine proteinase, useful as a cosmetic, medicament for treating

PT eczema, psoriasis, arthritis, and in the manufacture of the medicament

PT for treating, preventing pathogenic diseases involving receptor

PT mediated binding

XX XX

PS Disclosure; Page 5; 38pp; English.

CC The specification describes a fish serine proteinase. The proteinases

CC are useful as medicaments, for treating and preventing a disease in a

CC human or an animal such as pain, acute inflammation, chronic

CC inflammation, arthritis, inflamed joints, bursitis, osteoarthritis,

CC fibromyalgia, systemic lupus erythematosus, phlebitis, tendinitis, rash,

CC psoriasis, eczema, facial seboreic eczema, eczema of the hands,

CC face or neck, forsetin infections, athlete's foot, fistulae infections,

CC keloids, boils, warts and allergic itch, hemorrhoids, wounds, scars,

CC infections, wounds from burns, fungal infections and immunological and

CC autoimmune diseases. They are also useful for removing dead or peeling

CC skin from otherwise healthy skin, and for treating or preventing a

CC disease in which pathogenesis is caused by bacteria, virus, fungus,

CC parasite or a protozoan or a receptor mediated binding is involved.

CC The present sequence represents the amino terminal of bovine trypsin B

CC chain.

XX XX

Sequence 20 AA;

Query Match 100.0%; Score 112; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.7e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVNGEAVFGSPMPQVSLQD 20

DB 1 IVNGEAVFGSPMPQVSLQD 20

RESULT 2

AAAY99596

ID AAAY99596 standard; protein; 245 AA.

AC AAAY99596;

DT 13-SEP-2000 (first entry)

DE Bovine chymotrypsinogen A.

KW Bovine; plasminogen activator; cardiant; thrombolytic;

KW heart attack; stroke; blood clotting disorder.

XX Bos taurus.

XX W0200032759-AL.

XX 08-JUN-2000.

XX 06-MAY-1999; 99MO-US09991.

XX 02-DEC-1998; 98US-0110588.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PI Lin X, Zhang XC, Tang JJN;

XX WPI; 2000-422975/36.

XX Polypeptide with plasminogen activator activity useful as thrombolytic

PT agent for treating blood clot disorders e.g. heart attack, comprises 10

PT amino acid peptide fragment for recognition or activation of

PT plasminogen

XX XX

PS Disclosure; Page 35-36; 41pp; English.

XX The present sequence is bovine chymotrypsinogen. It was included in

CC a review of sequence homologues of several plasminogen activators.

CC plasminogen is the principal serine protease zymogen in the

CC extracellular fluids of vertebrates. Its active form, plasmin, is

CC implicated in pericellular proteolysis associated with a wide range of

CC physiological and pathological processes. Plasminogen expression is

CC regulated by plasminogen activators which hydrolyse a peptide bond in

CC plasminogen to convert it to plasmin or form tight binding complexes

CC with plasminogen to spontaneously convert it to plasmin. The sequence

CC homology analysis has identified a six amino acid peptide involved in

CC plasminogen activation. This peptide is particularly useful when inserted

CC between amino acid residues 644 and 645 of full length human

CC plasminogen. Novel plasminogen activators have been made based upon the

CC plasminogen activation/recognition site of plasminogen binding proteins.

CC The polypeptides are useful in preparing thrombolytic agents for treating

CC blood clotting disorders such as heart attack.

XX XX

Sequence 245 AA;

Query Match 100.0%; Score 112; DB 21; Length 245;

Best Local Similarity 100.0%; Pred. No. 2.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVNGEAVFGSPMPQVSLQD 20

DB 16 IVNGEAVFGSPMPQVSLQD 35

RESULT 3

AAAY33344

ID AAAY33344 standard; peptide; 20 AA.

AC AAAY33344;

DT 29-NOV-1999 (first entry)

XX Bovine chymotrypsin N-terminal fragment.

XX Multifunctional enzyme; treatment; acne; eczema; chymotrypsin; trypsin;

KW collagenase; elastase; exopeptidase; cell surface receptor; anti-acne;

KW antiseborrheic; dermatological; anti-eczema; adhesion molecule; ICAM-1;

KW CD54; ICAM-2; VCAM-1; CD4; CD8; CD28; CD29D; CD31; CD44; CD49; CD62L;

KW CD102; GM1 ceramide; krill; bovine.

XX Bos taurus.

XX US5958406-A.

XX 28-SEP-1999.

XX 08-FEB-1996; 96US-0600273.

XX 22-NOV-1994; 94US-0388501.

XX 08-FEB-1995; 95US-0385540.

XX 07-JUN-1995; 95US-0486820.

XX (PHAI-) PHAIRSON MEDICAL INC.

XX Lindblom R, Kay J, Franklin RL, De Faire JR;

XX WPI; 1999-561004/47.

XX Treating acne and eczema using a krill-derived multifunctional enzyme -

XX Disclosure; Column 21-22; 42pp; English.

CC This invention describes a novel method for treating acne and eczema

CC using a krill-derived multifunctional enzyme (I) which comprises 2 or

CC more of the activities of chymotrypsin, trypsin, collagenase, elastase

CC or exopeptidase and is reactive with cell surface receptors such as

CC proteins or glycoproteins. The product of the invention have

CC antiseborrheic, anti-acne, dermatological and anti-eczema activity.

CC (1) removes or inactivates cell surface receptors (proteins and
 CC glycoproteins) such as ICAM-1 (i.e. CD54), ICAM-2,
 CC VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD49, CD62L, CD102 and the
 CC asialo GM1 ceramide. This sequence represents the N-terminal fragment of
 CC bovine chymotrypsin which is used to describe the method of the
 CC invention.

XX Sequence 20 AA;

Query Match 97.3%; Score 109; DB 20; Length 20;
 Best Local Similarity 95.0%; Pred. No. 4.6e-09;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IYNGEEAVPGSMFQVSLQD 20
 |||||:|||||
 Db 1 IYNGEDAVPGSMFQVSLQD 20

RESULT 4
 AAY93935
 ID AAY93935 standard; peptide; 20 AA.
 XX AC AAY93935;

DT 03-OCT-2000 (first entry)
 XX DE N-terminal of a bovine chymotrypsin hydrolase enzyme.
 XX KW Transplantation rejection; hydrolase; graft versus host disease;
 KW cell surface adhesion molecule; immune reaction; inflammation; shock;
 KW tumour metastasis; autoimmune disease;
 KW Kill derived multifunctional enzyme.

XX Bos sp.
 XX PN W0200038708-A1.
 XX PD 06-JUL-2000.
 XX PF 23-DEC-1999; 99WO-US30818.
 XX PR 24-DEC-1998; 98US-0114147.
 XX PA (PHAI-) PHAIRSON MEDICAL INC.
 XX PI Franklin RL, St Pierre Y;
 XX WPI; 2000-452301/39.

XX Preventing or ameliorating transplantation rejection reactions using
 XX hydrolase enzymes -
 XX Disclosure; Page 26; 66pp; English.

CC The specification describes a method for preventing or ameliorating
 CC transplantation rejection reactions for transplantation of immune cells
 CC or other tissues. The method comprises providing a source of immune cells
 CC with a hydrolase or hydrolase mixture and administering the treated
 CC immune cells to a recipient animal. The hydrolase enzyme, at least
 CC preference for removing, destroying, inactivating or disabling at least
 CC one of CD4, CD8, CD25, CD28, ICAM-1 (CD54), CD152, an integrin, CD134,
 CC CD40 and CD80 in contrast to removing, destroying, inactivating or
 CC disabling TCR. The methods are useful for preventing graft versus host
 CC disease by using hydrolase enzymes to remove the cell surface adhesion
 CC molecules which are involved in triggering the immune reactions involved
 CC in the development of graft versus host disease.
 CC The methods are used for treating or preventing
 CC tumour metastasis, virus adhesion syndrome comprising inflammation, shock,
 CC cell-cell adhesion, viral adhesion disease, transplantation rejection
 CC reactions or microbial infections. The present sequence represents the
 CC N-terminal of a hydrolase, which may be used in the method of the
 CC invention.

XX Sequence 20 AA;

Query Match 97.3%; Score 109; DB 21; Length 20;
 Best Local Similarity 95.0%; Pred. No. 4.6e-09;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IYNGEEAVPGSMFQVSLQD 20
 |||||:|||||
 Db 1 IYNGEDAVPGSMFQVSLQD 20

RESULT 5
 AAE07941
 ID AAE07941 standard; peptide; 20 AA.
 XX AC AAE07941;

DT 01-NOV-2001 (first entry)
 XX KW N-terminal of bovine serine multifunctional hydrolase.
 XX KW Multifunctional hydrolase; rejection reaction; extra-corporeal;
 XX therapy; graft versus host disease; transplantation rejection;
 KW autoimmune disease; microbial infection; immune disorder; cytostatic;
 KW cystic fibrosis; chronic obstructive pulmonary disease; COPD;
 KW atherosclerosis; cancer; asthma; septic shock; toxic shock syndrome;
 KW conjunctivitis; reperfusion injury; pain; immunosuppressive;
 KW antibacterial; vasotropic; bovine.

XX Bos sp.

XX OS US5232088-B1.
 XX PN 15-MAY-2001.
 XX PD 24-DEC-1998; 98US-0220731.
 XX PR 08-FEB-1995; 95US-0385540.
 XX PR 07-JUN-1995; 95US-0482840.
 XX PR 08-FEB-1996; 96US-0600273.

XX (PHAI-) PHAIRSON MEDICAL INC.

XX Franklin RL, St Pierre Y;

XX WPI; 2001-450051/48.

XX Ameliorating, preventing or treating immune rejection reactions, such
 XX as graft versus host disease, autoimmune disease, asthma, cancer, by
 XX extra-corporeally treating donor tissue with hydrolase such as a
 XX protease -
 XX Disclosure; Column 15-16; 27pp; English.

CC The present invention relates to a method for ameliorating
 CC transplantation rejection reaction. The method comprises treating
 CC extra-corporeally donor tissue or donor selected immune cells with
 CC a rejection ameliorating effective amount of a hydrolase enzyme. The
 CC method is useful for ameliorating, treating or preventing disease.
 CC rejection reactions such as graft versus host diseases, organ or tissue
 CC transplantation rejection, autoimmune disease and associated conditions,
 CC microbial infection, immune disorder, cystic fibrosis, chronic
 CC obstructive pulmonary disease (COPD), atherosclerosis, cancer, asthma,
 CC septic shock, toxic shock syndrome, conjunctivitis, reperfusion injury
 CC or pain in humans. The present sequence is the N-terminal peptide
 CC of a bovine serine multifunctional hydrolase, used in the
 CC exemplification of the invention.

XX Sequence 20 AA;

Query Match 97.3%; Score 109; DB 22; Length 20;
 Best Local Similarity 95.0%; Pred. No. 4.6e-09;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSNPQVSLQD 20
 Db 1 IVNGEDAVPGSNPQVSLQD 20

RESULT 6
 ID AAB54191 standard; Protein; 146 AA.
 AC AAB54191;
 XX
 XX
 DT 09-MAR-2001 (first entry)
 XX
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:643.
 XX
 KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW neoplastic; immunomodulatory; relaxant; contraceptive; gynecological;
 KW antiinflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.

XX Homo sapiens.
 XX WO200129056-A1.
 XX
 XX 26-APR-2001.
 XX
 XX 20-OCT-2000; 2000WO-US29095.
 XX
 XX 20-OCT-1999; 99US-0421213.
 XX
 XX (UTAR-) UNIV ARKANSAS.
 XX
 XX O'Brien TJ, Tanimoto H;
 XX WPI; 2001-381031/40.
 XX
 XX Novel extracellular serine protease, termed tumor antigen-derived gene
 XX 15 protein overexpressed in carcinomas and DNA encoding it, for
 XX diagnosis, treatment, prevention of cancer, particularly breast,
 XX ovarian cancer -

XX Example 10; Fig 1, 130pp; English.
 XX
 XX The present invention relates to human tumor antigen-derived gene 15
 XX (TAGD-15) protein and coding sequence (see AAH23601 and AAB98500).
 XX TAGD-15 is an extracellular serine protease. It was found that TAGD-15 is
 XX over-expressed in ovarian tumors. TAGD-15 protein or its fragments of
 XX 9-20 residues that lack TAGD-15 protease activity are useful for
 XX vaccinating an individual against TAGD-15, having, suspected of having or
 XX at risk of getting cancer. Furthermore, the TAGD-15 gene can be used as a
 XX diagnostic or therapeutic target in cancer. The present sequence is used
 XX in a sequence homology alignment with the catalytic domain of TAGD-15.
 XX
 XX Sequence 231 AA;
 XX
 XX Query Match 97.3%; Score 109; DB 22; Length 231;
 XX Best Local Similarity 95.0%; Pred. No. 6.8e-08;
 XX Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSNPQVSLQD 20
 Db 2 IVNGEDAVPGSNPQVSLQD 21

RESULT 8
 ID AAU82738 standard; Protein; 263 AA.
 AC AAU82738;
 XX
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Amino acid sequence of novel human protease #37.

XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
 KW visual function; human immunodeficiency virus; HIV; non-viral infection;
 KW ocular disease; cytostatic; enzyme.
 XX Homo sapiens.
 OS
 XX W0200200860-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX
 PF
 XX 26-JUN-2001; 2001WO-US20171.
 XX
 PR 26-JUN-2000; 2000US-214047P.
 XX
 PA (SUVE-1) SUGEN INC.
 XX
 XX Plozman G. Whyte D. Sudarsanam S. Manning G. Caenepeel S;
 PI Charyczak G;
 XX
 XX WPI: 2002-139913/18.
 DR N-PSDB; ABK31780.
 XX
 XX Nucleic acids encoding novel human proteases, useful for useful for
 PT treating diseases and disorders such as cancer, immune-related
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
 PT inflammatory disorders.
 XX
 XX Claim 6; Fig 2W; 313pp; English.
 PS
 XX The present invention relates to the isolation of novel human
 CC proteases, and the nucleic acids encoding them. The sequences of
 CC the invention are useful for treating diseases and disorders such as
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
 CC (e.g. atherosclerosis and coronary thrombosis), brain or neuronal-associated
 CC diseases (e.g. rheumatoid arthritis and psoriasis), inflammatory
 CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
 CC peripheral nervous system disorders, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, neurological disorders, hypotension,
 CC hypertension, psychotic disorders, neurodegenerative disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease)
 CC The nucleic acids and polypeptides are also useful for
 CC infections caused by human immunodeficiency virus (HIV), and non-viral
 CC infections such as ocular disease (e.g. glaucoma) and macular
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
 CC the invention.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 97.1%; Score 109; DB 23; Length 263;
 Best Local Similarity 95.0%; Pred. No. 7.8e-08;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IVNGEEAVPGSNPWQVSLQD 20
 |||||
 Db 34 IVNGEDAVPGSNPWQVSLQD 53
 RESULT 9
 ID AAB11711 standard; Protein: 264 AA.
 XX
 XX AAB11711;
 XX
 XX 23-OCT-2000 (first entry)
 DT
 XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.
 DE
 XX BSSP5; serine protease; human; mBSSP5; mouse; mBSSP5; brain;
 KW

KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
 KW epilepsy; cancer; inflammation; infertility; pancreatitis;
 KW prostatic hypertrophy.
 OS
 XX Mus sp.
 PN W0200031243-A1.
 XX
 XX 02-JUN-2000.
 PD
 XX
 PF 19-NOV-1999; 99WO-JP06473.
 XX
 PR 20-NOV-1998; 98JP-0347806.
 XX
 PA (FUSO) FUSO PHARM IND LTD.
 XX
 XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
 DR WPI: 2000-400058/34.
 XX N-PSDB; AA61734.
 DR
 XX Serine proteases BSSP5, useful in detecting homologs, mutants and
 PT polymorphic variant markers for diagnosis of e.g. Alzheimer's
 PT disease, epilepsy, cancer and inflammation, using blood, urine,
 PT pancreas or other tissues.
 XX
 XX Claim 3; Page 55-56; 70pp; Japanese.
 PS
 XX The invention relates to novel serine proteases designated BSSP5
 CC (AAB11710-B11711), and to nucleic acids encoding them (AA61735-AA61734).
 CC The invention also relates to vectors and transfectants comprising BSSP5
 CC nucleic acids; transgenic animals in which the expression level of BSSP5
 CC can be varied; and an mBSSP5 knockout mouse. The invention additionally
 CC encompasses anti-BSSP5 antibodies and methods of production of such
 CC antibodies, methods of BSSP5 detection using the antibodies, and the
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain
 CC pancreatic conditions, e.g., pancreatitis. A method for detecting
 CC pancreatic conditions comprising measuring BSSP5 concentration in the blood or
 CC urine and also detecting nucleotides encoding BSSP5 were initially
 CC isolated in a human brain. Nucleotides encoding BSSP5 were initially
 CC (AA61744-AA61745) based on conserved regions of serine proteases. The
 CC BSSP5 serine proteases and nucleotides encoded by them are useful in
 CC detecting homologues, mutants and polymorphic variants in biological
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, pancreas
 CC and spleen) as diagnostic markers for conditions such as
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
 CC human BSSP5 (mBSSP5), and sequence AAB11711 represents murine BSSP5
 CC (mBSSP5).
 XX
 XX Sequence 264 AA;
 SQ
 Query Match 95.5%; Score 107; DB 21; Length 264;
 Best Local Similarity 95.0%; Pred. No. 1.5e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 IVNGEEAVPGSNPWQVSLQD 20
 |||||
 Db 34 IVNGEDAVPGSNPWQVSLQD 53
 RESULT 10
 ID AAB54077 standard; Protein: 192 AA.
 XX
 XX AAB54077;
 XX
 XX 09-MAR-2001 (first entry)
 DT
 XX Human pancreatic cancer antigen protein sequence SEQ ID NO:529.
 DE
 XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW

KW detection; diagnosis; identification; cytosolic; neuroprotective;
 KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW anti-inflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.
 XX Homo sapiens.
 XX WO200055320-A1.
 XX PD 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US05989.
 XX PR 12-MAR-1999; 99US-0124270.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 XX WPI; 2000-579444/54.
 XX N-PSDB; AAC98842.
 XX New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX Claim 11; Page 966; 1379pp; English.
 XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, the human pancreatic cancer antigens, given in AAB54008 to
 CC AAB54007, the human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,
 CC anticarcinogenic, cardiac and anti-inflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX SQ Sequence 192 AA;
 Query Match 86.6%; Score 97; DB 21; Length 192;
 Best Local Similarity 90.0%; Pred. NO. 36-06; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 IVNGEAVPGSWPQVSLQD 20
 Db 30 IVNGENAVLGSPWQVSLQD 49
 RESULT 11
 ID AAB11710 standard; Protein; 264 AA.
 XX AAB11710;
 XX AC AAB11710;
 XX DT 23-OCT-2000 (first entry)
 XX Human serine protease BSSP5 (hbSSP5) SEQ ID NO:2.

KW BSSP5, serine protease; human; hbSSP5; mouse; mBSSP5; brain;
 KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;
 KW epilepsy; cancer; inflammation; infertility; pancreatitis;
 KW prostatic hypertrophy.
 XX Homo sapiens.
 XX WO2000031243-A1.
 XX PD 02-JUN-2000.
 XX PF 19-NOV-1999; 99WO-JP06473.
 XX PR 20-NOV-1998; 98JP-0347806.
 XX PA (FUSO) FUSO PHARM IND LTD.
 XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;
 XX WPI; 2000-400058/34.
 XX N-PSDB; AAA61733.
 XX Serine proteases BSSP5, useful in detecting homologs, mutants and
 PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's
 PT disease, epilepsy, cancer and inflammation, using blood, urine,
 PT pancreas or other tissues -
 XX Claim 1; Page 51-52; 70pp; Japanese.
 XX The invention relates to novel serine proteases designated BSSP5
 CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).
 CC The invention also relates to vectors and transformants comprising BSSP5
 CC nucleic acids; transgenic animals in which the expression level of BSSP5
 CC can be varied; and BSSP5 antibodies and methods of production of such
 CC encompassing antibodies. BSSP5 detection using the antibodies, and the
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain
 CC medical conditions, e.g., pancreatitis. A method for detecting
 CC BSSP5 in a sample, e.g., blood, urine, brain, prostate gland, placenta, testis,
 CC paine, and a pancreatitis diagnostic agent containing an anti-BSSP5
 CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially
 CC isolated in a human brain cDNA library using degenerate PCR primers
 CC (AAA61744-A61745) based on conserved regions of serine proteases. The
 CC BSSP5 serine proteases and nucleotides encoding them are useful in
 CC detecting homologues, mutants and polymorphic variants in biological
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,
 CC pancreas and spleen) as diagnostic markers for conditions such as
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents
 CC human BSSP5 (hbSSP5), and sequence AAB11711 represents murine BSSP5
 CC (mBSSP5).
 XX SQ Sequence 264 AA;
 Query Match 86.6%; Score 97; DB 21; Length 264;
 Best Local Similarity 90.0%; Pred. NO. 4.2e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 IVNGEAVPGSWPQVSLQD 20
 Db 34 IVNGENAVLGSPWQVSLQD 53
 RESULT 12
 ID AAW02588 standard; peptide; 31 AA.
 XX AAW02588;
 XX AC AAW02588;
 XX DT 12-FEB-1997 (first entry)
 XX FX, porcine pancreas-derived factor N-terminal peptide.

XX Pig; bone; calcium; resorption inhibition; formation; osteoporosis;
 KW Paget's disease; fracture repair; bone defect; osteopetrosis;
 KW metastasis; osteoblast; osteoclast.
 XX
 OS Sus scrofa.
 XX
 PN MO9619501-A1.
 XX
 PD 27-JUN-1996.
 XX
 PF 19-DEC-1995; 95WO-US16826.
 XX
 XX 20-DEC-1994; 94US-0363092.
 XX
 XX (OSTE-) OSTEOSA INC.
 XX
 XX Izbicka E, Mundy GR, Yoneda T;
 XX WPI, 1996-309523/31.
 XX
 XX New isolated pancreatic-derived factor - which inhibits bone
 PT resorption and increases bone formation, used to develop prods. for
 PT diagnosis and therapy
 XX
 XX Claim 1; Fig 14; 53pp; English.

XX AAW02588 is the N-terminal portion of a factor derived from pig
 CC pancreatic tissue. The factor was designated PX. PX inhibits the
 CC action of osteoclasts and stimulates the action of osteoblasts. The
 CC hence PX inhibits bone resorption and stimulates bone formation. The
 CC new factor also stimulates the proliferation of human MG-63
 CC osteosarcoma cells (osteoblast-like cells). The factor is useful for
 CC treating bone loss, e.g. due to osteoporosis or Paget's disease. The
 CC factor may also be used to hasten bone fracture repair. Antibodies
 CC against osteoclasts of PX can be used to treat osteoporosis and
 CC osteoblastic metastasis
 XX
 SQ Sequence 31 AA;

Query Match 83.0%; Score 93; DB 17; Length 31;
 Best Local Similarity 84.2%; Pred. No. 1.5e-06;
 Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSNPMQVSLQ 19
 :||||:|||||
 Db 1 VVNGEDAVPYSPMOMVSLQ 19

RESULT 13
 ID AAF60061 standard; Protein; 242 AA.
 XX
 AC AAF60061;
 XX
 DT 06-SEP-1991 (first entry)
 XX
 DE Sequence of human pancreatic elastase IIIB.
 XX

XX Enzyme; serum lipoprotein metabolism.
 XX Homo sapiens.
 XX
 XX EP198645-A.
 XX
 PD 22-OCT-1986.
 XX
 PF 07-APR-1986; 86EP-0302557.
 XX
 PR 23-OCT-1985; 85JP-0216686.
 PR 05-APR-1985; 85JP-0072308.
 PR 27-APR-1985; 85JP-0091986.
 PR 26-JUL-1985; 85JP-0163964.

PR 02-DEC-1985; 85JP-0271128.
 XX
 XX (SANY) SANKYO KK.
 XX
 PT Takiguchi Y, Tani T, Kawashima I, Erukawa H, Ohmine T;
 PT Ohnami J;
 DR WPI; 1986-280300/43.
 DR N-PSDB; AAN60075.
 XX
 PT New genetically engineered human pancreatic elastase - obtcd.
 PT using hosts modified DNA coding for enzyme
 XX
 XX Claim 2; Page 8; 45pp; English.
 XX
 XX The genetically engineered prod. can eliminate the dependency on
 CC human pancreas supplies for the elastase, and avoids antibody
 CC formation and possibility of anaphylaxis using porcine elastase.
 XX
 SQ Sequence 242 AA;

Query Match 83.0%; Score 93; DB 7; Length 242;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;
 Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSNPMQVSLQ 19
 :||||:|||||
 Db 1 VVNGEDAVPYSPMOMVSLQ 19

RESULT 14
 ID AAR29621 standard; Protein; 253 AA.
 XX
 AC AAR29621;
 XX
 DT 30-APR-1993 (first entry)
 XX
 DE Porcine pancreatic elastase III.
 XX
 KW Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
 KW liver function improvers; ss.
 XX
 OS Sus scrofa domestica.
 XX
 XX JP04325090-A.
 XX
 XX 13-NOV-1992.
 XX
 PF 23-APR-1991; 91JP-0092069.
 PR 23-APR-1991; 91JP-0092069.
 XX
 XX (SANY) SANKYO CO LTD.
 XX
 XX WPI; 1992-428826/52.
 DR N-PSDB; AAQ31724.
 XX

PT Pig pancreas elastase protein - used in bile acid secretion
 PT promoters and liver function improvers
 XX
 XX Claim 1; Page 8; 10pp; Japanese.
 XX
 XX This sequence represents porcine pancreatic elastase III. It was
 CC expressed in E. coli YAZ1 using the expression vector pELE001.
 CC It can be used in bile acid secretion promoters and liver function
 CC improvers.
 XX
 SQ Sequence 253 AA;

Query Match 83.0%; Score 93; DB 13; Length 253;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;
 Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Wed Feb 12 11:59:41 2003

```
QY 1 IVNGEEAVPGSWPMQVSLQ 19
Db 12 VVNGEDAVPYSPWPMQVSLQ 30

RESULT 15
AAR84270
ID AAR84270 standard; peptide; 29 AA.
XX AAR84270;
XX 11-APR-1996 (first entry)
XX Serum calcium lowering factor N-terminal sequence.
XX Serum calcium lowering factor; rat; pancreas; calcium; mouse;
XX bone disease; osteoporosis; primary hyperthyroidism;
XX continuous hyperthyroidism; hypercalcaemia; malignant tumour.
XX Rattus rattus.
XX OS
XX JP07215997-A.
XX 15-AUG-1995.
XX 01-FEB-1994; 94JP-0027578.
XX 01-FEB-1994; 94JP-0027578.
XX (CHUS ) CHUGAI PHARM CO LTD.
XX WPI; 1995-317483/41.
XX Serum calcium lowering factor obtd. from rat pancreas - useful as
XX drug for treatment of various bone disease e.g. osteoporosis,
XX hyperthyroidism and hypercalcaemia.
XX Claim 2; Page 2; 6pp; Japanese.
XX This sequence represents the N-terminal sequence of a serum callum
XX lowering factor derived from rat pancreas. The factor has a decrease
XX molecular weight of 30000 measured by SDS-PAGE and causes a decrease
XX in serum levels of calcium in mice in a dose dependent manner. The
XX factor may be used for the prevention and treatment of various bone
XX diseases such as osteoporosis, primary and continuous hyperthyroidism
XX and hypercalcaemia accompanying malignant tumours.
XX Sequence 29 AA;
Query Match 80.4%; Score 90; DB 16; Length 29;
Best Local Similarity 78.9%; Pred. NO. 3.8e-06;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 IVNGEEAVPGSWPMQVSLQ 19
Db 1 VVNGEDAVPNSPMQVSLQ 19
```

Search completed: February 12, 2003, 10:22:26
Job time : 14.4328 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 / Search time 5.01493 Seconds
(without alignment)
383.393 Million cell updates/sec

Title: US-10-036-371-6

Perfect score: 112

Sequence: 1 IVNGEAVPGSWPQVSLQD 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR 73.*

1: piri.*

2: piri.*

3: piri.*

4: piri.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	245	1 KYBOA	chymotrypsin (EC 3
2	109	97.3	245	1 KYBOB	chymotrypsin (EC 3
3	109	97.3	263	2 A21195	chymotrypsin (EC 3
4	108	97.3	263	2 A21299	chymotrypsin (EC 3
5	108	96.4	263	1 KYRIB	chymotrypsin (EC 3
6	102	92.1	120	2 B61333	chymotrypsin (EC 3
7	98	87.5	246	2 A24773	chymotrypsin-like
8	98	87.5	246	2 A24773	chymotrypsin B - A
9	97	86.6	264	2 T31812	chymotrypsin-like
10	96	85.7	263	2 S47537	chymotrypsin (EC 3
11	93	83.0	31	2 B3257	pancreatic elastase
12	93	83.0	270	2 A29934	pancreatic elastase
13	88	78.6	270	2 B29934	pancreatic elastase
14	86	76.8	23	2 PU0036	serine proteinase
15	86	76.8	269	2 B26823	pancreatic elastase
16	84	75.0	271	1 ELRT2	pancreatic elastase
17	83	74.1	269	2 A26823	pancreatic elastase
18	83	74.1	271	2 A25528	pancreatic elastase
19	82	73.2	268	2 JQ4473	pancreatic elastase
20	82	72.3	269	2 C26823	pancreatic elastase
21	80	71.4	268	2 S68826	pancreatic elastase
22	80	71.4	268	2 S68826	pancreatic elastase
23	77	68.8	1528	2 T30338	pancreatic elastase
24	76	67.9	421	2 S29589	polyprotein - Afri
25	75	67.0	37	2 S02176	acrosin (EC 3.4.21
26	73	65.2	240	1 CP0R03	acrosin (EC 3.4.21
27	73	65.2	415	1 A34170	procarboxypeptidase
28	72	64.3	28	2 A61529	chymotrypsin (EC 3
29	70	62.5	1004	2 T30338	oviductin (EC 3.4.

30 69 61.6 20 2 A34817 collagenolytic pro
31 69 61.6 418 2 A52384 acrosin (EC 3.4.21
32 69 61.6 420 2 A52384 acrosin (EC 3.4.21
33 69 61.6 436 2 JX0172 acrosin (EC 3.4.21
34 69 61.6 638 1 KOWSPL plasma kallikrein
35 68 60.7 20 2 S33787 trypsinase (EC 3.4
36 68 60.7 25 2 A23698 trypsinase (EC 3.4
37 68 60.7 270 2 S56160 mast cell trypsinase
38 68 60.7 274 2 A45754 trypsinase (EC 3.4
39 68 60.7 275 2 A35863 trypsinase (EC 3.4
40 68 60.7 275 2 A35863 trypsinase (EC 3.4
41 68 60.7 275 2 C35863 trypsinase (EC 3.4
42 68 60.7 431 2 S47538 trypsinase (EC 3.4
43 67 59.8 230 2 I48685 mast cell proteinase
44 67 59.8 273 2 A47246 trypsinase (EC 3.4
45 67 59.8 274 2 JCA171 trypsinase (EC 3.4

ALIGNMENTS

RESULT 1

KYBOA

Chymotrypsin (EC 3.4.21.1) A precursor - bovine

N:Alternate names: chymotrypsinogen A

C:Species: Bos primigenius taurus (cattle)

C>Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 07-May-1999

C:Accession: A90235; A93158; S29650; A00952

B:Brown, J.R.; Hartley, B.S.

A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide

A:Reference number: A90235; PMID:67181721; PMID:5971783

A:Molecule type: protein

A:Residues: 1-101 'N' 103-245

R:Blow, D.M.; Birkoft, J.J.; Hartley, B.S.

Nature 211, 337-340, 1969

A:Title: Role of a buried acid group in the mechanism of action of chymotrypsin.

A:Reference number: A93158; PMID:6916266; PMID:5764436

A:Contents: annotation; revision to residue 102

R:McLoun, B.; Klueh, I.; Koelka, V.; Moravsek, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Sorm

Biochim. Biophys. Acta 130, 543-546, 1986

A:Title: Covalent structure of bovine chymotrypsinogen A.

A:Reference number: A90572; PMID:67183948; PMID:5972866

A:Molecule type: protein

A:Accession: A93158

A:Residues: 101 'N' 103-245 <ML>

A:Note: disulphide bonds determined

R:Citrizzola, F.; Ascenzi, P.; Bolognesi, M.; Bolognesi, E.; Sartti, P.; Schnel

Biochim. Biophys. Acta 1161, 201-208, 1993

A:Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catalytic

A:Reference number: S29650; PMID:9316038; PMID:8431470

A:Accession: S29650

A:Molecule type: protein

A:Residues: 1-12;16-27;149-160;181-200 <CUT>

R:Smillie, L.B.; Hartley, B.S.

Biochem. J. 101, 232-241, 1966

A:Title: Histidine sequences in the active centres of some 'serine' proteinases.

A:Reference number: A90236; PMID:67181723; PMID:5971785

R:Blow, D.M.; Blow, D.M.; Henderson, R.; Steltz, T.A.

Philos. Mag. 3, 8257, 67-76, 1970

A:Title: The structure of bovine alpha-chymotrypsin.

A:Reference number: A93754

A:Contents: annotation; X-ray crystallography

C:Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.

delta-chymotrypsin; further chymotryptic cleavage liberates the dipeptide His-Ile and As-

d An-148 directly from chymotrypsinogen, which leads to the degraded form neochymotryps

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F1-245/Product: chymotrypsinogen #status experimental <ZYM>

F1-13,16-146,149-245/Product: alpha-chymotrypsin #status experimental <NPT>

F:16-238/Domain: trypsin homology <TRY>
 F:1-122,42-58,136-201,168-182,191-220/Dисульфиде bonds; #status experimental
 F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 100.0%; Score 112; DB 1; Length 245;
 Best Local Similarity 100.0%; Pred. No. 3,5e-09; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0

QY 1 IVNGEAVPGSNPWQVSLQD 20
 |||||
 DB 16 IVNGEAVPGSNPWQVSLQD 35

RESULT 2
 KYRQB
 Chymotrypsin (EC 3.4.21.1) B precursor - bovine
 N:Alternate names: chymotrypsinogen B
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 18-Jul-1997
 C:Accession: A00953
 R:Synonyms: A00953
 R:Synonyms: L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.
 Nature 218, 343-346, 1968
 A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen
 F:16-238/Domain: trypsin homology <TRY>
 A:Reference number: A00953; MUID:68238908; PMID:5649671
 A:Accession: A00953
 A:Molecule type: protein
 A:Residues: 1-245 <SMI>
 C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acin
 C:Comment: The first activation cleavage, leading to pi-chymotrypsin B, occurs in the ac
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F:1-15/Domain: propeptide #status experimental <PRO>
 F:16-245/Product: chymotrypsin B #status experimental <MAT>
 F:16-238/Domain: trypsin homology <TRY>
 F:1-122,42-58,136-201,168-182,191-220/Dисульфиде bonds; #status experimental
 F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 97.3%; Score 109; DB 1; Length 245;
 Best Local Similarity 95.0%; Pred. No. 9.7e-09; Indels 0; Gaps 0;
 Matches 19; Conservative 1; Mismatches 0

QY 1 IVNGEAVPGSNPWQVSLQD 20
 |||||
 DB 16 IVNGEAVPGSNPWQVSLQD 35

RESULT 3
 A21195
 Chymotrypsin (EC 3.4.21.1) 2 precursor - dog
 C:Species: Canis lupus familiaris (dog)
 C>Date: 27-Jul-1993 #sequence_revision 27-Jul-1990 #text_change 22-Jun-1999
 C:Accession: A21195
 R:Synonyms: D. LaForge, K.S.; Luc, V.; Scheele, G.
 P:Finch, M.T.; AcadSci U.S.A. 80, 7486-7490, 1983
 A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence dete
 F:16-238/Domain: trypsin homology <TRY>
 A:Reference number: A21195; MUID:84170253; PMID:6584866
 A:Accession: A21195
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-263 <PIN>
 A:Cross-references: GB:K01173; NID:G163945; PIDN:AAA10841.1; PID:G163946
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:13-33/Domain: propeptide #status predicted <PRO>
 F:34-263/Product: chymotrypsin B #status predicted <MAT>
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 97.3%; Score 109; DB 2; Length 263;
 Best Local Similarity 95.0%; Pred. No. 1e-08; Indels 0; Gaps 0;
 Matches 19; Conservative 1; Mismatches 0

QY 1 IVNGEAVPGSNPWQVSLQD 20
 |||||
 DB 34 IVNGEAVPGSNPWQVSLQD 53

RESULT 4

A31299
 Chymotrypsin (EC 3.4.21.1) precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 22-Jun-1999
 C:Accession: A31299
 R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Matsub
 Biochem. Biophys. Res. Commun. 158, 569-575, 1989
 A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinoge
 A:Reference number: A31299; MUID:89134264; PMID:2917002
 A:Accession: A31299
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-263 <TM>
 A:Cross-references: GB:M24400; NID:G181189; PIDN:AAA52128.1; PID:G181190
 C:Genetics: GDB:CTRB1; CTRB
 A:Cross-references: GDB:119820; OMIM:118890
 A:Map position: 16q23.1-16q23.1
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 97.3%; Score 109; DB 2; Length 263;
 Best Local Similarity 95.0%; Pred. No. 1e-08; Indels 0; Gaps 0;
 Matches 19; Conservative 1; Mismatches 0

QY 1 IVNGEAVPGSNPWQVSLQD 20
 |||||
 DB 34 IVNGEAVPGSNPWQVSLQD 53

RESULT 5

KYRTB
 Chymotrypsin (EC 3.4.21.1) B precursor - rat
 N:Alternate names: chymotrypsinogen B
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
 C:Accession: A22658
 R:Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.
 J. Biol. Chem. 259, 14265-14270, 1984
 A:Title: Isolation and sequence of a rat chymotrypsin B gene.
 A:Reference number: A22658; MUID:85054881; PMID:6209274
 A:Accession: A22658
 A:Molecule type: DNA
 A:Residues: 1-263 <BEL>
 A:Cross-references: GB:K02298; NID:G203653; PIDN:AAA98732.1; PID:G203654
 C:Genetics: 18
 A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:13-33/Domain: propeptide #status predicted <PRO>
 F:34-263/Product: chymotrypsin B #status predicted <MAT>
 F:34-256/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 96.4%; Score 108; DB 1; Length 263;
 Best Local Similarity 90.0%; Pred. No. 1.5e-08; Indels 0; Gaps 0;
 Matches 18; Conservative 2; Mismatches 0

QY 1 IVNGEAVPGSNPWQVSLQD 20
 |||||
 DB 34 IVNGEAVPGSNPWQVSLQD 53

RESULT 6

B61333
 Chymotrypsin (EC 3.4.21.1) - edible frog (fragment)
 C:Species: Rana esculenta (edible frog)

C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
 C/Accession: B61333
 C/Species: Homo sapiens (man)
 C/Date: 16-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
 C/Title: Amino-terminal amino acid sequences and the evolution of frog (*Rana esculenta*)
 C/Reference number: A61333; MUID:80113255; PMID:6965480
 A/Molecule type: protein
 A/Residues: 1-148
 C/Superfamily: trypsin, trypsin homology
 C/Keywords: hydrolase; protein digestion; serine proteinase

Query Match 91.1%; Score 102; DB 2; Length 20;
 Best Local Similarity 90.0%; Pred. No. 6.6e-09; Mismatches 1; Indels 0; Gaps 0;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGEAVPGSNPQVSLQD 20
 Db 1 IVNGEAVPGSNPQVSLQD 20

RESULT 7

Chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)
 N/Alternate names: pancreatic elastase II (misidentification)
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
 C/Accession: A23473
 R/Vered, M.; Gertler, A.; Burshtein, Y.
 Int. J. Pept. Protein Res. 27, 183-190, 1986
 A/Reference number: A23473; MUID:86194934; PMID:3634756
 A/Accession: A23473
 A/Molecule type: protein
 A/Residues: 1-126 <VER>
 C/Superfamily: trypsin; trypsin homology
 C/Keywords: hydrolase; serine proteinase

Query Match 87.5%; Score 98; DB 2; Length 126;
 Best Local Similarity 70.0%; Pred. No. 2e-07; Mismatches 6; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVNGEAVPGSNPQVSLQD 20
 Db 12 IVNGEAVPGSNPQVSLQD 31

RESULT 8

Chymotrypsin B - Atlantic cod (fragments)
 C/Species: Gadus morhua (Atlantic cod)
 C/Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Aug-1998
 C/Accession: S72219
 R/Lieth-Larsen, R.; Asgerstrom, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.
 Biochim. Biophys. Acta 1287, 49-56, 1996
 A/Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.
 A/Reference number: S72219; MUID:96439045; PMID:8841380
 A/Accession: S72219
 A/Molecule type: Preliminary
 A/Residues: 1-14,15-244
 C/Superfamily: trypsin; trypsin homology
 F:15-237/Domain: trypsin homology <TRY>

Query Match 87.5%; Score 98; DB 2; Length 244;
 Best Local Similarity 94.7%; Pred. No. 4.1e-07; Mismatches 1; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGEAVPGSNPQVSLQ 19
 Db 15 IVNGEAVPHSNPQVSLQ 33

RESULT 9

Chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human
 C/Species: Homo sapiens (man)
 C/Date: 16-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
 C/Accession: B61333
 R/Larsen, P.; Solberg, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.
 Hum. Mol. Genet. 2, 1599-1595, 1993
 A/Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.
 A/Reference number: B31257; MUID:94095544; PMID:8268911
 A/Accession: B31257
 A/Molecule type: DNA
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Residues: 1-264 <RES>
 A/Cross-references: EMBL:X71874; NID:9406226; PIDN:CAA50710.1; PID:9406228
 C/Genetics:
 A/Gene: GDB:CTRL
 A/Cross-references: GDB:204061
 A/Map position: 16q22.1-16q22.1
 A/Map coordinates: 16/1; 52/3; 79/2; 106/3; 167/1; 211/3
 C/Superfamily: trypsin; trypsin homology
 C/Keywords: hydrolase; serine proteinase
 F:34-257/Domain: trypsin homology <TRY>
 F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 86.6%; Score 97; DB 2; Length 264;
 Best Local Similarity 90.0%; Pred. No. 6.3e-07; Mismatches 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVNGEAVPGSNPQVSLQD 20
 Db 34 IVNGEAVPGSNPQVSLQD 53

RESULT 10

Chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod
 C/Species: Gadus morhua (Atlantic cod)
 C/Date: 26-Dec-1997 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C/Accession: S47537
 R/Gudmundsdottir, A.; Oskarsson, S.; Ekin, A.E.; Craik, C.S.; Bjarnason, J.B.
 Biochim. Biophys. Acta 1219, 211-214, 1994
 A/Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.
 A/Reference number: S47537; MUID:94368860; PMID:8086467
 A/Accession: S47537
 A/Molecule type: mRNA
 A/Residues: 1-263 <GUD>
 A/Cross-references: EMBL:X78490; NID:9468750; PIDN:CAA55242.1; PID:9468751
 C/Superfamily: trypsin; trypsin homology
 C/Keywords: hydrolase; protein digestion; serine proteinase
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-263/Product: chymotrypsin #status predicted <MAT>
 F:75-237/Domain: trypsin homology <TRY>
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 85.7%; Score 96; DB 2; Length 263;
 Best Local Similarity 90.0%; Pred. No. 8.9e-07; Mismatches 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVNGEAVPGSNPQVSLQD 20
 Db 34 IVNGEAVPHSNPQVSLQD 53

RESULT 11

pancreatic elastase (EC 3.4.21.36) III - pig (fragment)
 N/Alternate names: proteinase E
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 30-Jun-1995
 C/Accession: B31257
 R/Larsen, P.; Solberg, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.
 Biochim. Biophys. Res. Commun. 163, 1191-1196, 1989
 A/Title: Generation of a submit III-like protein by autolysis of human and porcine prop

A:Reference number: A33257; MUID:8932022; PMID:2675835
 A:Accession: B33257
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-31 <AV>
 C:Superfamily: trypsin, trypsin homology
 C:Keywords: hydrolase; serine proteinase

Query Match 83.0%; Score 93; DB 2; Length 31;
 Best Local Similarity 84.2%; Pred. No. 2.3e-07;
 Matches 16; Conservative 1; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSWPMQVSLQ 19
 :|||||:|||||:|||||
 DB 12 VVNGEDAVPYSPWPMQVSLQ 30

RESULT 12
 B29934
 N:Alternate names: pancreatic elastase (EC 3.4.21.36) IIB precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Dec-1988 #sequence revision 15-Dec-1988 #text change 22-Jun-1999
 C:Accession: B29934; A90516; A27206; A33257; A28932; S04959; S04490
 R:Tani, T.; Ohnumi, J.; Mita, K.; Takiguchi, Y.
 J. Biol. Chem. 263, 1231-1239, 1988
 A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase
 A:Reference number: A92664; MUID:88087253; PMID:2826474
 A:Accession: B29934
 A:Molecule type: mRNA
 A:Residues: 1-270 <TAN>
 A:Cross-references: GB:M18692; NID:9607029; PID:AAA58454.1; PID:g182035
 R:Shen, W.; Fletcher, T.S.; Largman, C.
 Biochemistry 26, 3447-3452, 1987
 A:Title: Primary structure of human pancreatic protease B determined by sequence analysis
 A:Reference number: A90516; MUID:88000545; PMID:3477287
 A:Accession: A90516
 A:Molecule type: mRNA
 A:Residues: 'G', '5-63', 'G', '65-78', 'W', '80-118', 'G', '120-163', 'P', '165-270 <SHE>
 R:Fletcher, T.S.
 submitted to GenBank, August 1987
 A:Reference number: A94507
 A:Contents: revision to residue 119
 A:Accession: A27206
 A:Molecule type: DNA
 A:Residues: 'G', '5-63', 'G', '65-78', 'W', '80-163', 'P', '165-270 <FLE>
 R:Aviles, F.X.; Pascual, R.; Salva, M.; Bonicel, J.; Puigserver, A.
 Biochem. Biophys. Res. Commun. 163, 1191-1196, 1989
 A:Title: Generation of a substrate III-like protein by autocatalysis of human and porcine pro
 A:Reference number: A33257; MUID:8932022; PMID:2675835
 A:Accession: A33257
 A:Molecule type: protein
 A:Residues: 1-31 <AV>
 A:Keywords: hydrolase; serine proteinase
 A:Residues: 'X', '32-52', 'X', '54-55', 'XXX', '59-63 <GVY>
 R:Mouillard, M.; Kerfelec, B.; Mallet, B.; Chapus, C.
 FEBS Lett. 250, 166-170, 1989
 A:Title: Identification of a procarboxypeptidase A-truncated protease E binary complex
 A:Reference number: S04959; MUID:89325560; PMID:2753124
 A:Accession: S04959
 A:Molecule type: protein
 A:Residues: 31-50 <ND>
 R:Wendorf, P.; Geyer, R.; Sziegoleit, A.; Linder, D.
 FEBS Lett. 249, 275-278, 1989
 A:Title: Localization and characterization of the glycosylation site of human pancreatic
 A:Reference number: S04490; MUID:89289996; PMID:2737288
 A:Accession: S04490
 A:Molecule type: protein

A:Residues: 94-128;132-164 <WEN>
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: glycoprotein; hydrolase; serine proteinase; zymogen
 P:1-17/Domain: signal sequence #status predicted <SIG>
 P:18-28/Domain: activation peptide #status predicted <ACT>
 P:29-270/Product: (or 31-270) pancreatic elastase IIB #status predicted <MAT>
 P:29-263/Domain: trypsin homology <TRY>
 P:73-123,217/Active site: His, Asp, Ser #status predicted
 P:114/Binding site: carbohydrate (Asn) (covalent) #status experimental
 P:153/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 83.0%; Score 93; DB 2; Length 270;
 Best Local Similarity 84.2%; Pred. No. 2.5e-06;
 Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSWPMQVSLQ 19
 :|||||:|||||:|||||
 DB 29 VVNGEDAVPYSPWPMQVSLQ 47

RESULT 13
 A29934
 N:Alternate names: pancreatic elastase (EC 3.4.21.36) IIA precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Dec-1988 #sequence revision 15-Dec-1988 #text change 20-Jun-2000
 C:Accession: A29934; JX0045
 R:Tani, T.; Ohnumi, J.; Mita, K.; Takiguchi, Y.
 J. Biol. Chem. 263, 1231-1239, 1988
 A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase J
 A:Reference number: A92664; MUID:88087253; PMID:2826474
 A:Accession: A29934
 A:Molecule type: DNA
 A:Residues: 1-270 <TAN>
 A:Cross-references: GB:J03516
 R:Shimizu, Y.; Takemura, K.; Yoshida, H.; Sato, Y.; Iijima, H.; Shimada, Y.; Mikayama, T.
 J. Biochem. 104, 259-264, 1988
 A:Title: Molecular cloning of complementary DNA encoding one of the human pancreatic prot
 A:Reference number: JX0045; MUID:89034017; PMID:2460440
 A:Accession: JX0045
 A:Molecule type: mRNA
 A:Residues: 1-240, 'G', '242-270 <SHI>
 A:Cross-references: GB:D00306; NID:9220013; PID:BAA00212.1; PID:g220014
 C:Comment: This enzyme is an alanine-specific serine proteinase that has little elastoly
 C:Genetics:
 A:Introns: 15/1; 43/3; 76/2; 121/2; 167/1; 214/3; 265/3
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; pancreas; serine proteinase
 P:1-15/Domain: signal sequence #status predicted <SIG>
 P:16-28/Domain: propeptide #status predicted <PRO>
 P:29-270/Product: pancreatic elastase IIA #status experimental <MAT>
 P:29-263/Domain: trypsin homology <TRY>
 P:73,123,217/Active site: His, Asp, Ser #status predicted

Query Match 78.6%; Score 88; DB 2; Length 270;
 Best Local Similarity 78.9%; Pred. No. 1.4e-05;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSWPMQVSLQ 19
 :|||||:|||||:|||||
 DB 29 VVNGEDAVPYSPWPMQVSLQ 47

RESULT 14
 P00316
 N:Alternate names: trypsin-like proteinase bPTLP
 C:Species: Bos praedonatus laurus (cattle)
 C:Date: 28-Oct-1993 #sequence revision 28-Oct-1994 #text change 22-Apr-1995
 C:Accession: S04936; P00316
 R:Kishijima, A.; Sakai, K.; Edazawa, K.; Nagata, K.; Sasaki, Y.; Nagamune, H.; Matsuda, Y.
 submitted to JIBID, September 1994
 A:Description: Purification and characterization of a novel serine proteinase from bovine

```

A:Reference number: PU0036
A:Accession: PU0036
A:Molecule type: protein
A:Residues: 1-23 <RSU>
A:Experimental source: pancreas
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match
Best Local Similarity 76.8%; Score 86; DB 2; Length 23;
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSNPWQVLSQ 19
    :|||:|||||
Db 1 VVGGEAIPHSNPNQVLSQ 19

RESULT 15
B26823
pancreatic elastase II (EC 3.4.21.71) A precursor - human
C:Species: Homo sapiens (man)
C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 20-Jun-2000
R:Accession: B26823; A27432; A1431; S34491
DNA:Masuda, T.; Imai, T.; Shimoda, K.; Takiguchi, Y.
A:Title:Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed in the pancreas of the rat
A:Reference number: A90958; MUID:87217962; PMID:3646943
A:Accession: B26823
A:Molecule type: mRNA
A:Residues: 1-269 <KAM>
A:Cross-references: GB:M16652; NID:g182057; PIDN:AAAS2380.1; PID:g182058
R:Fletcher, T.S.; Shen, W.F.; Largman, C.
Biochemistry 26, 7256-7261, 1987
A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis
A:Reference number: A27432; MUID:88107669; PMID:3427074
A:Accession: A27432
A:Molecule type: mRNA
A:Residues: 1-269 <FJE>
A:Cross-references: GB:M16631; NID:g182022; PIDN:AAAS2374.1; PID:g182023
R:Shitaku, I.; Yoshida, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, T.
Biochemistry 12, 1555-1563, 1987
A:Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human F
A:Reference number: A1431; MUID:88198076; PMID:283436
A:Accession: A1431
A:Molecule type: mRNA
A:Residues: 1-201, 'V', 203-269 <SHI>
A:Cross-references: GB:D00236; NID:g219619; PIDN:BA00165.1; PID:g219620
A>Note: the authors translated the codon GTG for residue 202 as Cys
R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.
FEBS Lett. 261, 179-183, 1990
A:Title: Further studies on the human pancreatic binary complexes involving procarboxype
A:Reference number: S08253; MUID:90169111; PMID:2307232
A:Accession: S34491
A:Molecule type: protein
A:Residues: 'X', 18-50 <MOU>
A:Genet:GB:ELAI
A:Cross-references: GDB:119866; OMTM:130120
A:Map position: 12pter-12qter;
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; serine proteinase
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-28/Domain: propeptide #status predicted <PRO>
F:29-269/Product: pancreatic elastase IIA #status predicted <WAT>
F:29-262/Domain: trypsin homology <TRY>
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match
Best Local Similarity 76.8%; Score 86; DB 2; Length 269;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSNPWQVLSQ 19
    :|||:|||||

```

Db 29 VVGGEARPNNSWPNQVLSQ 47

Search completed: February 12, 2003, 10:28:52
Job time: 6.01493 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:04:45 ; Search time 2.68657 Seconds
(without alignments)
308.768 Million cell updates/sec

Title: US-10-036-371-6

Perfect score: 112

Sequence: 1 1VNGEEAVPGSWPQVSLQD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	245	1	CTRB_BOVIN
2	109	97.3	245	1	CTRB_BOVIN
3	109	97.3	263	1	CTRB_CANPA
4	109	97.3	263	1	CTRB_HUMAN
5	108	96.4	263	1	CTRB_HUMAN
6	98	87.5	245	1	CTRB_RAT
7	97	86.6	264	1	CTRB_GADMO
8	96	85.7	263	1	CTRB_GADMO
9	93	83.0	270	1	CTRB_HUMAN
10	90	80.4	268	1	ELCB_RAT
11	88	78.6	270	1	ELCB_HUMAN
12	85	76.8	269	1	ELCB_HUMAN
13	85	75.3	253	1	CAC3_BOVIN
14	84	75.0	292	1	ELCB_RAT
15	84	75.0	492	1	ELCB_HUMAN
16	83	74.1	269	1	ELCB_BOVIN
17	83	74.1	269	1	ELCB_HUMAN
18	81	72.3	269	1	ELCB_HUMAN
19	81	72.3	269	1	ELCB_HUMAN
20	80	71.4	268	1	ACRO_CAPHI
21	75	67.0	60	1	ACRO_PIG
22	75	67.0	437	1	TMS4_HUMAN
23	73	65.2	415	1	ACRO_PIG
24	73	65.2	418	1	HATT_HUMAN
25	72	64.3	490	1	TMS2_MOUSE
26	69	61.6	400	1	COGL_PARC
27	69	61.6	400	1	COGL_PARC
28	68	60.9	436	1	MCT6_MOUSE
29	68	60.9	436	1	MCT6_MOUSE
30	68	60.7	235	1	ELAS_GADMO
31	68	60.7	235	1	TRMT_HUMAN
32	68	60.7	270	1	MCT7_RAT
33	68	60.7	275	1	TRB1_HUMAN

RESULT 1

ID	CTRB_BOVIN	STANDARD	PRT	245 AA
AC	P00766;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsinogen A (EC 3.4.21.1).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	1:CB;_taxid=9913;			
RP	SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.			
RX	MEDLINE=67181721; PubMed=5971785;			
RA	Brown J.R., Hartley B.S.			
RT	Location of disulphide bridges by diagonal paper electrophoresis.			
RT	The disulphide bridges of bovine chymotrypsinogen A."			
RL	Biochem. J. 101:214-228(1966).			
RN	[2]			
RP	REVISION TO 102.			
RX	MEDLINE=69106266; PubMed=5764436;			
RA	Blow D.M., Birktoft J.J., Hartley B.S.;			
RT	"Role of a buried acid group in the mechanism of action of			
RT	chymotrypsin."			
RL	Nature 221:337-340(1969).			
RP	PRELIMINARY SEQUENCE.			
RA	Hartley B.S.			
RT	"Amino acid...sequence of bovine chymotrypsinogen-A."			
RL	Nature 201:1284-1287(1964).			
RN	[4]			
RP	SEQUENCE, AND DISULFIDE BONDS			
RX	MEDLINE=67183948; PubMed=5972866;			
RA	Mellon B., Kluh I., Kostka V., Moravsek L., Prusik Z., Vanacek J.			
RT	"Covalent structure of bovine chymotrypsinogen A."			
RL	Biochim. Biophys. Acta 130:543-546(1966).			
RN	[5]			
RP	ACTIVE SITE.			
RX	MEDLINE=67181723; PubMed=5971785;			
RA	Smillie L.B., Hartley B.S.;			
RT	"Histidine sequences in the active centres of some 'serine'			
RT	proteases."			
RL	Biochem. J. 101:232-241(1966).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY			
RX	MEDLINE=72035052; PubMed=439050;			
RA	Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;			
RT	"I. Serine proteinases. The structure of alpha-chymotrypsin."			
RL	Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).			
RN	[7]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.			
RX	MEDLINE=7017557; PubMed=5442169;			
RA	Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;			

P20231 homo sapien
P15157 homo sapien
Q9u152 homo sapien
Q9u152 oryctolagus
Q9u152 mus musculus
P50343 mus musculus
P21845 mus musculus
P21845 mus musculus
Q9u152 homo sapien
Q9u152 mus musculus
Q9u152 mus musculus
P12886 ovis aries
Q16651 homo sapien

ALIGNMENTS

RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation.";

RL Biochemistry 9:11997-2009(1970).

RN (8)

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.

RX MEDLINE=82078042; PubMed=6914359;

RA Cohen G.H., Silverton E.W., Davies D.R.;

RA "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.

RT Comparison with other pancreatic serine proteases.";

RL J. Mol. Biol. 148:449-479(1981).

RN (9)

RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.

RX MEDLINE=86011575; PubMed=4046030;

RA Tsukada H., Blow D.M.;

RA "Structure of alpha-chymotrypsin refined at 1.68-A resolution.";

RL J. Mol. Biol. 170:111(1983).

CC -I- ENZYME ACTIVITY: Preferential cleavage: Tyr|-Xaa, Trp|-Xaa,

CC -I- DATA: Xaa=1|-Xaa.

CC -I- SUBCELLULAR LOCATION: Extracellular.

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 51.

CC -I- DATABASE: NAME=Worthington enzyme manual;

CC WWW=http://www.worthington-biochem.com/manual/c/CHV.html".

DR PIR: R00952; KYBOA.

DR PDB: 2CGA; 15-APR-90.

DR PDB: 2CHA; 31-MAY-84.

DR PDB: 4CHA; 29-OCT-85.

DR PDB: 5CHA; 16-OCT-87.

DR PDB: 6CHA; 16-OCT-87.

DR PDB: 1CHG; 27-JAN-84.

DR PDB: 1CHO; 16-JUL-88.

DR PDB: 2GCH; 31-MAY-84.

DR PDB: 3GCH; 15-OCT-92.

DR PDB: 4GCH; 15-OCT-90.

DR PDB: 5GCH; 15-OCT-90.

DR PDB: 6GCH; 15-OCT-90.

DR PDB: 7GCH; 15-OCT-91.

DR PDB: 8GCH; 15-JUL-93.

DR PDB: 1GCT; 15-OCT-91.

DR PDB: 2GCT; 15-OCT-91.

DR PDB: 3GCT; 15-OCT-91.

DR PDB: 1ACB; 31-OCT-93.

DR PDB: 1GMC; 31-OCT-93.

DR PDB: 1GMD; 31-OCT-93.

DR PDB: 1CGI; 30-APR-94.

DR PDB: 1CGJ; 30-APR-94.

DR PDB: 1GCD; 22-JUN-94.

DR PDB: 1GHA; 22-JUN-94.

DR PDB: 1GHB; 22-JUN-94.

DR PDB: 1GHI; 30-SEP-94.

DR PDB: 2GMI; 01-NOV-94.

DR PDB: 1INT; 17-AUG-96.

DR PDB: 1AB9; 20-AUG-97.

DR PDB: 1AFO; 17-SEP-97.

DR PDB: 1CAO; 23-JUL-97.

DR PDB: 1CBW; 23-JUL-97.

DR PDB: 1VGC; 12-NOV-97.

DR PDB: 2VGC; 12-NOV-97.

DR PDB: 3VGC; 12-NOV-97.

DR PDB: 1VGC; 12-NOV-97.

DR PDB: 1VGC; 12-NOV-97.

DR PDB: 1VGC; 12-NOV-97.

DR PDB: 1VGC; 12-NOV-97.

DR MEMOPS: S01-001-

DR INTERPRO: IPR001314; Chymotrypsin.

DR Pfam: PF00089; trypsin; 1

DR Pfam: PF00089; trypsin; 1

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART: SM00320; Tryp_SPC; 1.

DR PROSITE: PS00240; TRYPSIN_DOM; 1.

DR PROSITE: PS00134; TRYPSIN_HIS; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;

SW 3D-structure. 1 13 CHYMOTRYPSIN A, A CHAIN.

FT CHAIN 16 146 CHYMOTRYPSIN A, B CHAIN.

CHYMOTRYPSIN A, C CHAIN.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.

FT CHAIN 149 245
FT ACT_SITE 57 57
FT ACT_SITE 102 102
FT ACT_SITE 195 195
FT DISULFID 1 122
FT DISULFID 42 58
FT DISULFID 136 201
FT DISULFID 168 182
FT DISULFID 191 220
FT HELIX 12 15
FT STRAND 20 21
FT TURN 24 25
FT TURN 28 29
FT STRAND 30 34
FT STRAND 40 46
FT STRAND 51 54
FT TURN 57 58
FT TURN 62 63
FT STRAND 65 68
FT TURN 69 69
FT STRAND 72 72
FT TURN 73 74
FT STRAND 81 90
FT TURN 92 93
FT STRAND 95 95
FT TURN 96 99
FT STRAND 100 100
FT TURN 101 101
FT STRAND 104 108
FT STRAND 122 122
FT TURN 126 127
FT TURN 132 133
FT STRAND 135 140
FT TURN 146 148
FT STRAND 154 154
FT STRAND 156 163
FT STRAND 165 166
FT TURN 169 172
FT HELIX 173 175
FT STRAND 180 181
FT TURN 185 186
FT TURN 195 196
FT STRAND 198 203
FT TURN 204 205
FT STRAND 206 216
FT TURN 218 219
FT TURN 222 223
FT STRAND 225 230
FT STRAND 231 233
FT TURN 234 234
FT HELIX 235 243
SQ SEQUENCE 245 AA; 25666 MW; 91A9F28E2F3E3142 CRC64;

Query Match 100.0%; Score 112; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVNGEEAVPGSGWPMQVSLQD 20
|||||
DB 16 IVNGEEAVPGSGWPMQVSLQD 35
|||||

RESULT 2

CTRB_BOVIN STANDARD; PRT; 245 AA.
ID_CTRB_BOVIN
AC P00767;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;


```

RP SEQUENCE OF 1-12 AND 16-31.
RC TISSUE=Pyloric caeca;
RX MEDLINE=92111252; PubMed=1764912;
RA Agreissron B., Bjarnason J.B.;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
  Gadus morhua". Comparison with bovine chymotrypsin.";
RI Cadmus morhua. Physiol. 1998;327-335(1991).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
CC Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR HSP; P00766; 1CHG.
DR MEROPS; S01.152; -.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; P00722; CHYMOTRYPSIN.
DR SMART; SM00202; TRYPSIN_DOM; 1.
DR PROSITE; PS00440; TRYPSIN_SPC; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 16 243 CHYMOTRYPSIN B, A CHAIN.
FT ACT_SITE 57 57 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 1 121 BY SIMILARITY.
FT DISULFID 42 58 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 167 182 BY SIMILARITY.
FT DISULFID 191 220 BY SIMILARITY.
FT CONFLICT 9 11 QVT -> VIS (IN REF. 2).
FT CONFLICT 26 26 S -> T (IN REF. 2).
FT CONFLICT 28 29 PW -> Y (IN REF. 2).
SQ SEQUENCE 245 AA; 26260 MW; 74FE0D425517AB02 CRC64;

Query Match 87.5%; Score 98; DB 1; Length 245;
Best Local Similarity 94.7%; Pred. No. 4.2e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LVNGEAVPGSWPMQVSLQ 19
Db 16 LVNGEAVPGSWPMQVSLQ 34

RESULT 7
CTRL_HUMAN STANDARD; PRT; 264 AA.
AC P40133;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin-like protease CTRL-1 precursor (RC 3.4.21.-).
GN CTRL OR CTRL1.
OS Homo sapiens (Human).
OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eumetaria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TAXID=9606;
[1]_TAXID=9606;
RN SEQUENCE FROM N.A.
RX MEDLINE=9403544; PubMed=8268911;
RA Larsen P., Solheim J., Kristensen T., Kolaro A.B., Prydz H.;
RT "A tight cluster of five unrelated human genes on chromosome
  16q22.1.";
RL Hum. Mol. Genet. 2:1589-1595(1993).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.eib-sib.ch/announce/
  or send an email to license@sib-sib.ch)
CC CC
CC EMBL; X71874; CAA50710.1; -
DR EMBL; X71877; CAA50711.1; -
DR HSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR Genew; HGNC:2524; CTRL.
DR MIM; 118889; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; P00722; CHYMOTRYPSIN.
DR SMART; SM00222; TRYPSIN_DOM; 1.
DR PROSITE; PS00440; TRYPSIN_SPC; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 33 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 34 264 CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 19 141 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 155 220 BY SIMILARITY.
FT DISULFID 187 201 BY SIMILARITY.
FT DISULFID 210 239 BY SIMILARITY.
SQ SEQUENCE 264 AA; 28002 MW; 3F629F02FA6DDFB4 CRC64;

Query Match 86.6%; Score 97; DB 1; Length 264;
Best Local Similarity 90.0%; Pred. No. 6.2e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LVNGEAVPGSWPMQVSLQ 20
Db 34 LVNGEAVPGSWPMQVSLQ 53

RESULT 8
CTRL_GADMO STANDARD; PRT; 263 AA.
AC P47796; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin (Atlantic cod) (RC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TAXID=8049;
[1]_TAXID=8049;
RN SEQUENCE FROM N.A.
RX TISSUE=Pyloric caeca;
RX MEDLINE=9436860; PubMed=8086467;
RA Gudmundsdottir A., Oskarsen S., Eakin A.E., Craik C.S.,
RA Bjarnason J.B.;
RT "Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";
RL Biochim. Biophys. Acta 1219:211-214(1994).
[2]_TAXID=8049;
RN SEQUENCE OF 19-30 AND 34-49.
RC TISSUE=Pyloric caeca;
RX Agreissron B., Bjarnason J.B.;
RX MEDLINE=92111252; PubMed=1764912;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
  Gadus morhua". Comparison with bovine chymotrypsin.";
RI Comp. Biochem. Physiol. 1998;327-335(1991).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,
CC Phe-|-Xaa, Leu-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.

```

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-aib.ch/announce/
CC or send an email to license@isb-aib.ch).
CC -----
CC EMBL: X78490; CAA55242.1; -.
CC HSSP: P00766; ICHG.
CC MEROPS: S01.152; -.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYP_SPC; 1.
CC PROSITE: PS50240; TRYPIN_DOM; 1.
CC PROSITE: PS00134; TRYPIN_HIS; 1.
CC PROSITE: PS00135; TRYPIN_SER; 1.
CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
CC SIGNAL 1 18
CC CHAIN 19 263 CHYMOTRYPSIN A.
CC ACT SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 19 140 BY SIMILARITY.
CC DISULFID 60 76 BY SIMILARITY.
CC DISULFID 154 219 BY SIMILARITY.
CC DISULFID 186 200 BY SIMILARITY.
CC DISULFID 209 238 BY SIMILARITY.
CC CONFLICT 21 21 R -> S (IN RSP. 2).
CC CONFLICT 25 25 S -> Q (IN RSP. 2).
CC CONFLICT 29 29 T -> S (IN RSP. 2).
CC CONFLICT 44 44 S -> T (IN RSP. 2).
CC CONFLICT 46 46 S -> Y (IN RSP. 2).
CC SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;
CC QUERY MATCH 85.7%; Score 96; DB 1; Length 263;
CC Best Local Similarity 90.0%; Pred. No. 8.6e-07;
CC Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 IYNGEAVPGSWQVSLQD 20
DB 34 IYNGEAVPGSWQVSLQD 53
RESULT 9
E13B HUMAN
AD E13B HUMAN STANDARD; PRT; 270 AA.
AC P08951; P11423;
DT 01-NOV-1988 (Rel. 09, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase IIIB precursor (EC 3.4.21.70) (protease B).
CS ELA3B human (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_taxid=9606;
ON 1
RE SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Medline=8808725; PubMed=2826474;
RA Tani T., Ohsumi J., Mita K., Takiguchi Y.;
RT "Identification of a novel class of elastase isozyme, human
RT pancreatic elastase III, by cDNA and genomic gene cloning.";
RN J. Biol. Chem. 263:1231-1239(1988).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RC Strausberg R.;

```

```

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 4-270 FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=88000545; PubMed=3477287;
RA Shen W., Fletcher T.S., Largman C.;
RT "Primary structure of human pancreatic protease B determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:3447-3452(1987).
RN [4]
RN SEQUENCE OF 31-50.
RC TISSUE=Pancreas;
RX MEDLINE=89345560; PubMed=2753124;
RA Moulard M., Kerfelec B., Mallet B., Chapus C.;
RT "Identification of a procarboxypeptidase A-truncated protease B
RT binary complex in human pancreatic juice.";
RL FEBS Lett. 250:166-170(1989).
RN [5]
RN SEQUENCE OF 94-164, AND CARBOHYDRATE-LINKAGE SITES ASN-114.
RC TISSUE=Pancreas;
RX MEDLINE=8289996; PubMed=2737288;
RA Wendorff P., Geyer R., Szegolet A., Linder D.;
RT "Localization and characterization of the glycosylation site of human
RT pancreatic elastase 1.";
RL FEBS Lett. 249:275-278(1989).
CC -1- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
CC LITTLE ELASTOTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Ala-Xaa. Does not
CC hydrolyse elastin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -1- CAUTION: Was originally (ref.5) thought to be elastase 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-aib.ch/announce/
CC or send an email to license@isb-aib.ch).
CC -----
CC EMBL: M16630; AAA36482.1; -.
CC EMBL: BC005216; AAH05216.1; -.
CC EMBL: M18692; AA58454.1; -.
CC PIR: B29934; B29934.
CC PIR: A27206; A27206.
CC PIR: S04999; S04999.
CC PIR: S04490; S04490.
CC HSSP: P05805; IFON.
CC MEROPS: S01.205; -.
CC GlycoSuiteDB: P08861; -.
CC SWISS-2DPAGE: P08861; HUMAN.
CC Genew: HGNC:15945; ELA3B.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYP_SPC; 1.
CC PROSITE: PS50240; TRYPIN_DOM; 1.
CC PROSITE: PS00134; TRYPIN_HIS; 1.
CC PROSITE: PS00135; TRYPIN_SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
CC SIGNAL 1 15
CC CHAIN 16 28 ACTIVATION PEPTIDE (POTENTIAL).
CC PROPEP 16 28 ELASTASE IIIB.
CC ACT SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 58 74 BY SIMILARITY.
CC DISULFID 117 120 PROBABLE.
CC DISULFID 157 223 BY SIMILARITY.
CC DISULFID 188 204 BY SIMILARITY.
CC DISULFID 213 244 BY SIMILARITY.
CC CARBOHYD 114 114 N-LINKED (GLCNAC. .).

```

```

FT CONFLICT 4 4 /FTID=CAR 000212.
FT CONFLICT 64 64 R -> G (IN REF. 3).
FT CONFLICT 79 79 A -> G (IN REF. 3).
FT CONFLICT 129 131 W -> R (IN REF. 1).
FT CONFLICT 164 164 MISSING (IN REF. 1).
FT CONFLICT 164 164 R -> P (IN REF. 3).
SQ SEQUENCE 270 AA; 29293 MW; B14HEUAAU369SAFE CRC64;

Query Match 81.0%; Score 93; DB 1; Length 270;
Best Local Similarity 84.2%; Pred No. 2 4e-06;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IVNGEEAVPGSNPMQVSLQ 19
DB 29 VVNGEDAVPSNPMQVSLQ 47

RESULT 10
CLCR_RAT STANDARD; PRT; 268 AA.
AC P55091; 063188;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DS Caldecrin precursor (EC 3.4.21.2) (Chymotrypsin C) (Serum calcium-
decreasing factor).
GN CTSC
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=96107178; PubMed=8530454;
RA Tomomura A., Tomomura M., Fukushima T., Akiyama M., Kubota N.,
RA Kunaki K., Nishii Y., Nolkura T., Saheki T.;
RT "Molecular cloning and expression of serum calcium-decreasing factor
(caldecrin).";
RL J. Biol. Chem. 270:30315-30321(1995).
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RX MEDLINE=92035053; PubMed=1537555;
RA Kagi J., Wieser J., Mueller-Hill B.;
RT "Identification of cDNAs encoding two novel rat pancreatic serine
proteases.";
RL Gene 110:181-187(1992).
RN CHARACTERIZATION.
RX MEDLINE=98207038; PubMed=9538241;
RA Yoshino-Yasuda T., Kobayashi K., Akiyama M., Itoh H., Tomomura A.,
RA Saheki T.;
RT "Caldecrin is a novel-type serine protease expressed in pancreas, but
its homologue, elastase IV, is an artifact during cloning derived
from caldecrin gene.";
RL J. Biochem. 123:546-554(1998).
RC -!- FUNCTION: HAS CHYMOTRYPSIN-TYPE PROTEASE ACTIVITY AND HYPOCALCEMIC
ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Cytoplasm; Xaa, Gln-I-Xaa, Asn-I-Xaa.
CC -!- TISSUE SPECIFICITY: Pancreas.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE ELASTASE IV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; S80379; AAB35830.1; -.
DR EMBL; X59014; CAA4753.1; -.
DR HSSP; P00766; 1CHG.
DR MEROPS; S01.157; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR PIR; PF00089; trypsin; 1.
DR PRINTS; PM00722; CHYMOTRYPSIN.
DR SMART; SM0020; Tryp_Spc; 1.
DR PROSITE; PS0040; TRYPSIN_DOM. 1.
DR PROSITE; PS00134; TRYPSIN_RES. 1.
DR PROSITE; PS00135; TRYPSIN_SEQ. 1.
KW Hydrolase, Serine protease, Glucoprotein, Zymogen; Signal.
PT SIGNAL 1 16 POTENTIAL.
PT PROPEP 17 29 ACTIVATION PEPTIDE.
PT CHAIN 30 268 CALDECRIN.
FT ACT_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 17 141 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 42 42 P -> A (IN REF. 2).
FT CONFLICT 96 120 REGSVYAEVDIIYHKNRLPLN -> AEAPCTLRWPTS
ISMRSGDSSGCG (IN REF. 2).
SQ SEQUENCE 268 AA; 29374 MW; 33B57AF3AD0F8983 CRC64;

Query Match 80.4%; Score 90; DB 1; Length 268;
Best Local Similarity 78.9%; Pred No. 6 4e-06;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IVNGEEAVPGSNPMQVSLQ 19
DB 30 VVNGEDAVPSNPMQVSLQ 48

RESULT 11
ELJA_HUMAN STANDARD; PRT; 270 AA.
AC P09093; Q9BRW4;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2001 (Rel. 41, Last annotation update)
DE Elapase AIIA precursor (EC 3.4.21.70) (Protease E).
GN ELA3
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RX MEDLINE=88087253; PubMed=2826474;
RA Tani T., Ohsumi J., Mita K., Takiguchi Y.;
RT "Identification of a novel class of elastase isozyme, human
pancreatic elastase III, by cDNA and genomic gene cloning.";
RL J. Biol. Chem. 263:12311-12319(1988).
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Pancreas;
RX MEDLINE=90907000; PubMed=2460440;
RA Yoshino-Yasuda T., Kobayashi K., Yoshida H., Sato Y., Iijima H.,
RA Shinada Y., Miyayama T., Ohtsuka T., Ikeda N., Ishida A., Tamai Y.,
RA Matuki S., Tanaka J., Ikenaga H., Ogawa M.;
RT "Molecular cloning of complementary DNA encoding one of the human
pancreatic protease E isozymes.";
RL J. Biochem. 104:259-264(1988).
RN SEQUENCE FROM N.A.
RP

```

```

RC TISSUE=Prostate;
RA Strauberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY
CC LITTLE ELASTOLYTIC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala|-Xaa. Does not
CC hydrolyse elastin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/announcement/
CC or send an email to license@ebi.ac.uk)
CC
CC -----
CC EMBL; M18700; AAA66350.1; -.
CC EMBL; M18693; AAA66350.1; JOINED.
CC EMBL; M18694; AAA66350.1; JOINED.
CC EMBL; M18695; AAA66350.1; JOINED.
CC EMBL; M18696; AAA66350.1; JOINED.
CC EMBL; M18697; AAA66350.1; JOINED.
CC EMBL; M18698; AAA66350.1; JOINED.
CC EMBL; M18699; AAA66350.1; JOINED.
CC EMBL; M18700; AAA66350.1; JOINED.
CC EMBL; BC005918; AAH05918.1; -.
CC EMBL; PIR; A29934; A29934.
CC HSP; P05805; IFON.
CC MEROPS; S01.154; -.
CC Genew; HGNC:15944; ELA2A.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRY_SP; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KX Hydroxylase; Serine protease; Zymogen; Signal; Glycoprotein.
CC FT PROPEP 1 23
CC CHAIN 26 278
CC ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 58 74 BY SIMILARITY.
CC DISULFID 117 120 PROBABLE.
CC DISULFID 157 203 BY SIMILARITY.
CC DISULFID 188 204 BY SIMILARITY.
CC DISULFID 213 244 BY SIMILARITY.
CC CARBOHYD 114 114 N-LINKED (GLCNAC...) (POTENTIAL).
CC CONFLICT 63 63
CC CONFLICT 106 106 MISSING (IN REF. 1).
CC CONFLICT 174 174 K -> E (IN REF. 3).
CC SEQUENCE 270 AA; 29474 MW; 576DDB255AA4118C CRC64;
CC
CC Query Match 78.6%; Score 88; DB 1; Length 270;
CC Best Local Similarity 78.9%; Pred. No. 1.3e-05;
CC Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CC
CC Oy 1 IVNGEAVPGSWPMQVSLQ 19
CC Db 29 VVHGDAVPGSWPMQVSLQ 47
CC
CC RESULT 12
CC EL2A_HUMAN STANDARD; PRT; 269 AA.
CC ID EL2A_HUMAN
CC AC P08217; Q14243;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-AUG-1988 (Rel. 08, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87217962; PubMed=3464943;
RA Kawahina I., Tani T., Shimoda K., Takiguchi Y.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RN mRNAs are expressed in human pancreas.";
RN DNA 6:163-172(1987).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=88107669; PubMed=3427074;
RA Fletcher T.S., Shen W.P., Largman C.;
RT "Primary structure of human pancreatic elastase 2 determined by
RN sequence analysis of the cloned mRNA.";
RN Biochemistry 26:7256-7261(1987).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=88198076; PubMed=283346;
RA Shirasu Y., Yoshida H., Matsuoka S., Takemura K., Ikeda N., Shimada Y.,
RA Ozawa T., Mikayama T., Iijima H., Sato Y., Tanai Y.,
RA Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RN encoding human pancreatic elastase 2.";
RN J. Biochem. 102:1555-1563(1987).
RN [4]
RN SEQUENCE FROM N.A.
RA Thomas D.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RX Strauberg R.;
RX TISSUE=Pancreas;
RA Strauberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa
CC and the Xaa hydrolyzes elastin.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/announcement/
CC or send an email to license@ebi.ac.uk)
CC
CC -----
CC EMBL; M16631; AAAS2174.1; -.
CC EMBL; M16652; AAAS2380.1; -.
CC EMBL; D00236; BAA00165.1; -.
CC EMBL; AL512883; CAC42421.1; -.
CC EMBL; BC007031; AAH07031.1; -.
CC PIR; A27432; A27432.
CC PIR; B26823; B26823.
CC HSP; P00772; IELG.
CC MEROPS; S01.155; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRY_SP; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KX Hydroxylase; Serine protease; Zymogen; Signal.
CC FT SIGNAL 1 16
CC PROPEP 17 28 ACTIVATION PEPTIDE.

```



```

FT CHAIN 29 269 ELASTASE 2A
FT DISULFID 58 74 BY SIMILARITY
FT DISULFID 155 222 BY SIMILARITY
FT DISULFID 186 202 BY SIMILARITY
FT DISULFID 212 243 BY SIMILARITY
FT ACT SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY)
FT ACT SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY)
FT ACT SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY)
FT CONFLICT 202 202 C -> V (IN REF. 3)
SQ SEQUENCE 269 AA; 28888 MW; A2E05143BFF4987C CRC64;

Query Match 76.9%; Score 86; DB 1; Length 269;
Best Local Similarity 78.9%; Pred. No. 3.4e-05;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IYNGEAVPGSNWQVSLQ 19
DB 29 VVGGEARPNWQVSLQ 47

RESULT 13
ID _CAC3 BOVIN STANDARD; PRT; 253 AA.
AC P05805;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Proproteinase E precursor (Procarboxypeptidase A complex component
DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).
OS Bos laurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-25.
RX MEDLINE=9109520; PubMed=2269366;
RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,
RA "Analysis of proproteinase E in bovine procarboxypeptidase A ternary
RA complex gives rise to subunit III.",
RL FEBS Lett. 277:37-41(1990).
RN [2]
RP SEQUENCE OF 14-253, AND DISULFIDE BONDS.
RX MEDLINE=86220190; PubMed=3519215;
RA Venot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G.,
RA "Amino acid sequence and disulfide bridges of subunit III, a
RA defective endopeptidase present in the bovine pancreatic 6 S
RA procarboxypeptidase A complex.",
RL Eur. J. Biochem. 157:91-99(1986).
RN [3]
RX X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RP MEDLINE=9422022; PubMed=8168476;
RX Pignol D., Gaboriau C., Michon T., Kerfelec B., Chapus C.,
RX Fontecilla-Camps J.C.;
RP "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a
RP defective endopeptidase present in the bovine pancreatic 6 S
RP EMBO J 13:1763-1771(1994) cited synonym E.5;
CC -1- FUNCTION: DEFECTIVE ELASTASE-LIKE SERINE PROTEASE. DOES NOT SEEM
CC TO HAVE A PROTEASE ACTIVITY. ITS LIKELY FUNCTION IS TO PROTECT
CC PROCARBOXYPEPTIDASE A AGAINST DENATURATION IN THE ACIDIC
CC ENVIRONMENT OF THE RUMINANT DUODENUM.
CC -1- SUBUNIT: HETEROTRIMER OF SUBUNIT III; CARBOXYPEPTIDASE A AND
CC CHYMOTRYPSINOGEN C.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: PANCREAS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR PIR: A25065; CPROA3.
DR PDB: 1F0N; 14-OCT-96.
DR MEROPS: S01.983; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00089; trypsin; 1.

```

```

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00250; TRYPTIC DOM. 1.
DR PROSITE; PS00130; TRYPSIN_DOM. 1.
DR PROSITE; PS00134; TRYPSIN_SER. 1.
DR PROSITE; PS00135; TRYPSIN_SER. 1.
KW Serine protease homolog; Pancreas; Digestion; 3D-structure.
FT PROPEP 1 11 ACTIVATION PEPTIDE.
FT CHAIN 12 253 PROPROTEINASE E.
FT DISULFID 41 57
FT DISULFID 100 103
FT DISULFID 140 206
FT DISULFID 171 187
FT DISULFID 196 227
SQ SEQUENCE 253 AA; 27337 MW; 24663724D9AE409C CRC64;

Query Match 75.9%; Score 85; DB 1; Length 253;
Best Local Similarity 78.9%; Pred. No. 3.2e-05;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IYNGEAVPGSNWQVSLQ 19
DB 12 VVGGEAVPYSNWQVSLQ 30

RESULT 14
ID _EL2 RAT STANDARD; PRT; 271 AA.
AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
DE ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9102967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic propeptidases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.",
RL Biochemistry 21:1453-1463(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054882; PubMed=6094548;
RA Swift G.H., Craik C.S., Stary S.J., Quinto C., Lahaie R.G.,
RA Rutter W.J., Macdonald R.J.;
RT "Structure of the two related elastase genes expressed in the rat
RT pancreas.",
RL J. Biol. Chem. 259:14271-14278(1984).
CC -1- FUNCTION: ACTS UPON ELASTIN.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-|Xaa
CC -1- SUBCELLULAR LOCATION: cytosol.
CC -1- SUBCELLULAR LOCATION: elastin.
CC -1- TISSUE SPECIFICITY: PANCREAS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; V01233; CAA24543.1; -.
DR EMBL; L00124; AAA98780.1; -.
DR EMBL; L00118; AAA98780.1; JOINED.
DR EMBL; L00119; AAA98780.1; JOINED.

```

```

DR EMBL; L00120; AAA98780.1; JOINED.
DR EMBL; L00121; AAA98780.1; JOINED.
DR EMBL; L00122; AAA98780.1; JOINED.
DR EMBL; L00123; AAA98780.1; JOINED.
DR PIR; A00961; ELRT2.
DR HSSP; P00772; 1ELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Zymogen; signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 271
FT ACT_SITE 75 75
FT ACT_SITE 123 123
FT ACT_SITE 218 218
FT DISULFID 60 76
FT DISULFID 157 224
FT DISULFID 188 204
FT DISULFID 214 245
FT DISULFID 214 245
SQ SEQUENCE 271 AA; 28885 MW; 125C783B857B71E3 CRC64;

Query Match 75.08; Score 84; DB 1; Length 271;
Best Local Similarity 73.74; Pred. No. 4.7e-05;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGQVPGSGKPAQVSLQ 19
DB 31 VVGQGRFNSWPAQVSLQ 49

RESULT 15
ID TM82 HUMAN STANDARD; PRT; 492 AA.
AC O1533; Q9BX11.
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 2 precursor (RC 3.4.21.-).
GN TMPSR2 OR PRSS10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1] _
SEQUENCE FROM N.A.
RX MEDLINE=97468144; PubMed=9325052;
RA Paoloni-Giacobino A., Chen H., Peitsch M.C., Rossier C.,
RA Antonarakis S.E.;
RA "Cloning of the TMPSR2 gene, which encodes a novel serine protease
RT with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3.";
RL Genomics 44:309-320(1997).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=21309069; PubMed=1141763;
RA Teng D.H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.;
RA "Mutation analyses of 268 candidate genes in human tumor cell lines.";
RL Genomics 74:352-364(2001).
[3]
SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=21139112; PubMed=11245484;
RA Afar D.E.H., Vivanco I., Hubert R.S., Kuo J., Chen B., Saffran D.C.,
RA Raicano A.B., Jakubovits A.;
RA "Catalytic cleavage of the androgen-regulated TMPSR2 protease results
RT in its secretion by prostate and prostate cancer epithelia.";
RL Cancer Res. 61:1686-1692(2001).
[4]

TISSUE SPECIFICITY.
MEDLINE=21104370; PubMed=11169526;
RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vibko P.T.;
RT "Expression of transmembrane serine protease TMPSR2 in mouse and
human tissues.";
RL J. Pathol. 193:134-140(2001).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ACTIVATED BY
CC CLEAVAGE AND SECRETED.
CC -1- TISSUE SPECIFICITY: EXPRESSED STRONGLY IN SMALL INTESTINE. ALSO
CC EXPRESSED IN PROSTATE, COLON, STOMACH, AND SALIVARY GLAND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/announcement/
CC or send an email to license@ebi.ac.uk).
-----
EMBL; U75329; AAC51784.1; -.
EMBL; AF123453; AAD37117.1; -.
EMBL; AF270487; AAK29280.1; -.
HSSP; P00763; LDPO.
DR MEROPS; S01.247; -.
DR GENE; HGNC:11876; TMPSR2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00132; LDLA; 1.
DR SMART; SM00202; SR; 1_SPC; 1.
DR PROSITE; PS00240; TRYD_SPC; 1.
DR PROSITE; PS00069; LDLRA_2; 1.
DR PROSITE; PS00068; LDLRA_1; 1.
DR PROSITE; PS00068; SRCR_1; FALSE_NEG.
DR PROSITE; PS00237; SRCR_2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Transmembrane; Signal-anchor; Zymogen;
Polyomorphism.
FT CHAIN 1 255
FT CHAIN 256 492
FT CHAIN 492 492
FT DOMAIN 1 84
FT TRANSMEM 85 105
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT LDL-RECEPTOR CLASS A.
FT DOMAIN 106 492
FT DOMAIN 112 149
FT DOMAIN 150 242
FT DOMAIN 256 492
FT ACT_SITE 296 296
FT ACT_SITE 345 345
FT ACT_SITE 441 441
FT SITE 255 256
FT DISULFID 113 126
FT DISULFID 120 139
FT DISULFID 132 148
FT DISULFID 172 231
FT DISULFID 185 244
FT DISULFID 281 297
FT DISULFID 410 426
FT DISULFID 437 465
FT CARBOHYD 213 213
FT CARBOHYD 249 249

```

FT VARIANT 449 449 K -> N (IN DBSNP:1056602).
FT MITAGEN 255 255 /FTID=VAR_011692;
FT MITAGEN 441 441 R->Q: LOSS OF CLEAVAGE.
FT MITAGEN 160 160 S->A: LOSS OF ACTIVITY.
FT CONFLICT 242 242 M -> V (IN REF. 3).
FT CONFLICT 329 329 I -> L (IN REF. 1).
FT CONFLICT 489 491 E -> Q (IN REF. 1).
FT CONFLICT 489 491 RAD -> KAN (IN REF. 1).
SQ SEQUENCE 492 AA; 53891 MN; CAB44FD17A9076B CRC64;

Query Match 75.0%; Score 84; DB 1; Length 492;
Best Local Similarity 77.8%; Pred. No. 8.4e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IYNGEAVPGSWFQVSL 18
||| | | | | | | | | |
Db 256 IYGESALPQWFWQVSL 273

Search completed: February 12, 2003, 10:23:18
Job time : 3.68657 secs

THIS PAGE BLANK (USPTO)

DR EMBL; AK003060; BAB22539.1; ..
 DR HSSP; P00766; 1GCT.
 DR MEROPS; S01.152; ..
 DR MGD; MGI:1913723; 2200008D09R1K.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27821 MW; 2620A2AFBASD04D CRC64;

Query Match 96.4%; Score 108; DB 11; Length 263;
 Best Local Similarity 90.0%; Pred. No. 2.5e-08;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IYNGEAVPGSNWQVSLQD 20
 Db 34 IYNGEAVPGSNWQVSLQD 53

RESULT 2

ID Q9DX8 PRELIMINARY; PRT; 263 AA.
 AC Q9DX8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 2200008D09R1K protein.
 DE 2200008D09R1K.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Saito K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito I., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirli L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Offield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
 TRYPSIN FAMILY.
 DR EMBL; AK007566; BAB25112.1; ..
 DR HSSP; P00766; 1GCT.
 DR MEROPS; S01.152; ..
 DR MGD; MGI:1913723; 2200008D09R1K.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

Query Match 96.4%; Score 108; DB 11; Length 263;
 Best Local Similarity 90.0%; Pred. No. 2.5e-08;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IYNGEAVPGSNWQVSLQD 20
 Db 34 IYNGEAVPGSNWQVSLQD 53

RESULT 3

ID Q9CR35 PRELIMINARY; PRT; 263 AA.
 AC Q9CR35;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 2200008D09R1K protein.
 DE 2200008D09R1K.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Saito K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito I., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirli L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Offield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
 TRYPSIN FAMILY.
 DR EMBL; AK008927; BAB25971.1; ..
 DR EMBL; AK003079; BAB22563.1; ..
 DR EMBL; AK007765; BAB25241.1; ..
 DR EMBL; AK007815; BAB25280.1; ..
 DR EMBL; AK008729; BAB25861.1; ..
 DR EMBL; AK008888; BAB25954.1; ..
 DR HSSP; P00766; 1GCT.
 DR MEROPS; S01.152; ..
 DR MGD; MGI:1913723; 2200008D09R1K.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

```

Query Match          96.4%; Score 108; DB 11; Length 263;
Best Local Similarity 90.0%; Pred. No. 2.9e+08;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSNPMQVSLQD 20
Db 34 IVNGEATPGSNPMQVSLQD 53

RESULT 4
Q9D960
ID Q9D960 PRELIMINARY; PRT; 264 AA.
AC Q9D960
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1810004D15RIK protein.
DE 1810004D15RIK.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX Medline=21085660; PubMed=11217951;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi X.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyone P., Marchionni L., Mashima J., Mazzarelli J., Nomaerts P.,
RA Nardone P., Ring B., Ringwald M., Rodriguez S., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR ENBL; AK007333; BAB24967.1; -.
DR HSSP; P00766; 4CHA.
DR MGD; MG1:88558; Ctrl.
DR MEROPS; S01:256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; I.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match          95.5%; Score 107; DB 11; Length 264;
Best Local Similarity 95.0%; Pred. No. 3.6e+08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSNPMQVSLQD 20
Db 34 IVNGEAVPGSNPMQVSLQD 53

RESULT 6
Q9QZ28
ID Q9QZ28 PRELIMINARY; PRT; 264 AA.
AC Q9QZ28
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

Q9D7P8
ID Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1810004D15RIK protein.
DE 1810004D15RIK.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX Medline=21085660; PubMed=11217951;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi X.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyone P., Marchionni L., Mashima J., Mazzarelli J., Nomaerts P.,
RA Nardone P., Ring B., Ringwald M., Rodriguez S., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR ENBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MGD; MG1:88558; Ctrl.
DR MEROPS; S01:256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; I.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match          95.5%; Score 107; DB 11; Length 264;
Best Local Similarity 95.0%; Pred. No. 3.6e+08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSNPMQVSLQD 20
Db 34 IVNGEAVPGSNPMQVSLQD 53

RESULT 6
Q9QZ28
ID Q9QZ28 PRELIMINARY; PRT; 264 AA.
AC Q9QZ28
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Sogane Y., Mitsu S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -.
DR HSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyse; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; 9F95DD10FD3500E CRC64;

Query Match 95.5%; Score 107; DB 11; Length 264;
Best Local Similarity 95.0%; Pred. No. 3.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IVNGEAVPGSWPQVSLQD 20
DB 34 IVNGEAVPGSWPQVSLQD 53

RESULT 7
ID OSER05 PRELIMINARY; PRT; 264 AA.
AC OSER05;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Chymopasin (Chymotrypsin A CTRA-1).
GN CTRL OR CTRA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Mitsu S., Yamaguchi N.;
RT "Molecular cloning of mouse chymopasin.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB014238; BAB20275.1; -.
DR HSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MG1.86558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;

```

```

Query Match 95.5%; Score 107; DB 11; Length 264;
Best Local Similarity 95.0%; Pred. No. 3.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IVNGEAVPGSWPQVSLQD 20
DB 34 IVNGEAVPGSWPQVSLQD 53

RESULT 8
ID OSW7Q3 PRELIMINARY; PRT; 260 AA.
AC OSW7Q3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Chymotrypsinogen 2.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029754; BAA82366.1; -.
DR HSP; P00766; 1CHG.
DR MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 90.2%; Score 101; DB 13; Length 260;
Best Local Similarity 90.0%; Pred. No. 2.8e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 IVNGEAVPGSWPQVSLQD 20
DB 31 IVNGEALPHSWPQVSLQD 50

RESULT 9
ID OSW7Q4 PRELIMINARY; PRT; 261 AA.
AC OSW7Q4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;

```


RT *Japanese flounder mRNA for chymotrypsinogen 1.*;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.

DR EMBL: AB029753; BA82345.1; -.
 DR HSSP: S00766; 1CHG.
 DR MEROPS: S01.154; 1UPO.

DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR001254; Ser. protease_Try.

DR Pfam: PF00089; trypsin_1.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART: SM0020; Tryp_SPC; 1.

DR PROSITE: PS02440; TRYPSIN_DOM; 1.

DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 261 AA; 28184 MW; D7090AD65395B7D CRC64;

Query Match 87.5%; Score 98; DB 13; Length 261;

Best Local Similarity 90.0%; Pred. No. 7.8e-07;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVNGEEAVPGSWPMQVSLQ 20

DB 32 LVNGEAVPGSWPMQVSLQ 51

RESULT 10

Q9PMQ6

ID Q9PMQ6 PRELIMINARY; PRT; 263 AA.

AC Q9PMQ6

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Chymotrypsin B precursor (BC 3.4.21.1).

CHV8

GN Gadus morhua (Atlantic cod).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Anthromorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.

OX NCBI_TaxId=8049;

RN 111_TaxId=8049;

RP SEQUENCE FROM N.A.

RC TISSUE=PYLORIC CAECA.

RX MEDLINE:20464334; PubMed1011764;

RA Spilliaert R., Gudmundsdottir A.

RL "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B.*";

RL Microb. Comp. Genomics 5:41-50(2000).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL: AJ242521; CAB43766.1; -.
 DR HSSP: P00766; 1CHG.
 DR MEROPS: S01.152; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. Protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM0020; Tryp_SPC; 1.
 DR PROSITE: PS02440; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 263 CHYMOTRYPSIN B
 FT SEQUENCE 263 AA; 28175 MW; 8F61B18A14E5E7C CRC64;

Query Match 87.5%; Score 98; DB 13; Length 263;

Best Local Similarity 94.7%; Pred. No. 7.9e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVNGEEAVPGSWPMQVSLQ 19

DB 32 LVNGEAVPGSWPMQVSLQ 50

RESULT 11

Q9D7T9

ID Q9D7T9 PRELIMINARY; PRT; 269 AA.

AC Q9D7T9

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE 2310074F01R1K protein.

GN ELA3B OR 2310074F01R1K.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RX NCBI_TaxId=10090;

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=STOMACH;

RX MEDLINE:2108560; PubMed121781;

RA Kawai J., Shingawa K., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Amanaka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kawakawa S., Sato K.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casanova H.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Watanabe S., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hwang S.-B., Aizawa K., Hasegawa Y., Kawaji H., Kohsaki S.,

RT "Functional Annotation of a full-length mouse cDNA collection.*";

RT Nature 409:685-690(2001).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL: AK008858; BAB25932.1; -.
 DR HSSP: P05805; 1PON.
 DR MEROPS: S01.154; -.
 DR MGD: MGI:1915118; Elab3.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM0020; Tryp_SPC; 1.
 DR PROSITE: PS02440; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 269 AA; 28977 MW; 9F43F769DDB2A7CF CRC64;

Query Match 86.6%; Score 97; DB 11; Length 269;

Best Local Similarity 89.5%; Pred. No. 1.1e-06;

Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVNGEEAVPGSWPMQVSLQ 19

DB 28 VVNGEEAVPHSPMQVSLQ 46

RESULT 12

Q9C052

ID Q9C052 PRELIMINARY; PRT; 269 AA.

AC Q9C052

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DE 2310074F01R1K protein.

GN ELA3B OR 2310074F01RIK.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP
RC
RC MEDLINE=2108560; PubMed=1217851;
RA Katsukawa J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Nakajima T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aikawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Kleeschmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Niki K., Tomita M., Wagner L.,
RA Schriml L.M., Staubli P., Suzuki R., Tonita M., Wagner L.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Machionni L., Mashima J., Mazzarelli J., Momaberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK010149; BAB26734.1; -;
DR EMBL; AK009129; BAB26032.1; -;
DR HSSP; P05805; 1FON.
DR MEROPS; S01.154; -;
DR MGD; MG1.191518; Elab3b.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR KW Hydrolyase; Serine protease.
DR KW Hydrolyase; Serine protease.
SQ SEQUENCE 269 AA; 28904 MW; C543F76957B2A7CE CRC64;
Query Match 86.68; Score 97; DB 11; Length 269;
Best Local Similarity 89.54; Pred. No. 1.1e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 INVGEAVPGSNPWQVSLQ 19
Db 28 VVNGEEAVPGSNPWQVSLQ 46
RESULT 13
QPPSP2
AC QPPSP2 PRELIMINARY; PRT; 40 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsin C, pI isoform (Fragments).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN
RN
RP
RX MEDLINE=95066525; PubMed=7976036;
RA Tluscik F., Polanowski A., Guyonnet V., Long P.L., Travis J.;

RT "Affinity purification of chicken pancreas proteinases and their N-
RT terminal amino-acid sequences";
RL Acta Biochim. Pol. 41:174-177(1994).
KW HSSP; P00766; 1CHG.
KW Hydrolyase; Serine protease.
FT NON_TER 1 1
FT NON_TER 16 17
FT NON_TER 40 40
FT NON_TER 40 40
SQ SEQUENCE 40 AA; 4376 MW; 348CFFBFC8C8003D CRC64;
Query Match 84.84; Score 95; DB 13; Length 40;
Best Local Similarity 89.54; Pred. No. 1e-07; 1; Indels 0; Gaps 0;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 INVGEAVPGSNPWQVSLQ 19
Db 17 IVHGEAVPGSNPWQVSLQ 35
RESULT 14
O19023 PRELIMINARY; PRT; 257 AA.
ID O19023
AC O19023 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-1998 (TrEMBLrel. 20, Last annotation update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DR Elacase (EC 3.4.21.36) (Fragment).
DR ELAC.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
OX (1)
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Szigoleit A.;
RT "Pancreatic elastase from rhesus monkey.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AJ000067; CAA03899.1; -;
DR HSSP; P05805; 1FON.
DR MEROPS; S01.154; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR KW Hydrolyase; Serine protease.
DR KW Hydrolyase; Serine protease.
FT NON_TER 1 1
FT NON_TER 257 257
FT NON_TER 257 257
SQ SEQUENCE 257 AA; 27687 MW; 4D443DB67233D8DC CRC64;
Query Match 83.04; Score 93; DB 6; Length 257;
Best Local Similarity 84.24; Pred. No. 4.1e-06;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 1 INVGEAVPGSNPWQVSLQ 19
Db 16 VVNGEAVPGSNPWQVSLQ 34
RESULT 15
Q660L8 PRELIMINARY; PRT; 270 AA.
ID Q660L8
AC Q660L8
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DR

```

DE Similar to elastase 3, pancreatic (protease E).
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_taxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC08383; AAH08383.1; -.
DR MEROPS; S01.154; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 270 AA; 29446 MW; 946DDDA694A102E CRC64;

Query Match          78.6%; Score 88; DB 4; Length 270;
Best Local Similarity 78.9%; Pred. No. 2.6e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IVNGEAVPGSMFQVSLQ 19
   :|:|:|:|:|:|:|:|:|
Db 29 VVHGDAVPSMFQVSLQ 47

```

Search completed: February 12, 2003, 10:27:26
Job time : 15.1493 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:18:40 ; Search time 4.29851 Seconds
(with alignments)
136.898 Million cell updates/sec

Title: US-10-036-371-6

Perfect score: 112

Sequence: 1 1VNGEAVPGSMQVSLQD 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A.COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B.COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A.COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B.COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCFUS.COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	228	1	US-08-278-091-10
2	112	100.0	228	1	US-08-483-859-10
3	112	100.0	228	1	US-08-472-173-10
4	112	100.0	228	2	US-08-487-167-10
5	112	100.0	228	2	US-08-482-816-10
6	112	100.0	228	2	US-08-226-149-10
7	112	100.0	228	2	US-08-801-499-10
8	112	100.0	228	2	US-08-615-271-10
9	112	100.0	228	3	US-09-074-660-10
10	112	100.0	228	3	US-09-074-659-10
11	112	100.0	228	3	US-09-106-468-10
12	112	100.0	228	4	US-08-106-468A-10
13	112	100.0	228	4	US-08-386-567-10
14	109	97.3	20	2	US-08-600-273A-16
15	109	97.3	20	3	US-08-486-820-16
16	109	97.3	20	4	US-09-220-731-16
17	109	97.3	229	2	US-08-557-146-13
18	109	97.3	229	2	US-09-154-344-13
19	109	97.3	230	4	US-08-944-483-62
20	109	97.3	231	4	US-09-027-337-6
21	109	97.3	231	4	US-09-644-600-6
22	93	83.0	31	5	US-09-16826-1
23	93	83.0	32	5	US-09-16826-2
24	93	83.0	32	5	US-09-16826-2
25	93	83.0	242	4	US-08-944-483-57
26	88	76.6	242	4	US-08-944-483-58
27	86	76.6	241	4	US-08-944-483-59

Sequence 1, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 8, Appl
Sequence 9, Appl
Sequence 10, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 15, Appl
Sequence 16, Appl
Sequence 17, Appl
Sequence 18, Appl
Sequence 19, Appl
Sequence 20, Appl
Sequence 21, Appl
Sequence 22, Appl
Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-278-091-10
; Sequence 10, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: FANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: CHONG, Raymond P.
; APPLICANT: KLEIN, Mark H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA: Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-278-091-10

Query Match 100.0%; Score 112; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 116-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 1VNGEAVPGSMQVSLQD 20
|||||

```

Db      1 1VNGEEAVPGSWPQVSLQD 20

RESULT 2
US-08-483-859-10
; Sequence 10, Application US/08483859
; Patent No. 5656436
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,859
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS.vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; FAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-173-10

Query Match      100.0%; Score 112; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 1VNGEEAVPGSWPQVSLQD 20
Db      1 1VNGEEAVPGSWPQVSLQD 20

RESULT 4
US-08-487-167-10
; Sequence 10, Application US/08487167
; Patent No. 5869302
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,167
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-487-167-10

Query Match 100.0%; Score 112; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVPGSNPWVSLQD 20
DB 1 INVGEAVPGSNPWVSLQD 20

RESULT 5
US-08-482-816-10
; Sequence 10, Application US/08482816
; Patent No. 5935573
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sam & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,816
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-482-816-10

Query Match 100.0%; Score 112; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVPGSNPWVSLQD 20
DB 1 INVGEAVPGSNPWVSLQD 20

RESULT 6
US-08-296-149-10
; Sequence 10, Application US/08296149
; Patent No. 593297
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sam & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,149
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-296-149-10

Query Match 100.0%; Score 112; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVPGSNPWVSLQD 20
DB 1 INVGEAVPGSNPWVSLQD 20

RESULT 7

```

```

US-08-601-499-10
; Sequence 10, Application US/08801499
; Patent No. 5962430
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 51m & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601,499
; FILING DATE: 20-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,816
; FILING DATE: 07-JUN-1995
; PRACTICE INFORMATION:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-601-499-10

Query Match 100.0%; Score 112; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPQVSLQD 20
Db 1 IVNGEAVPGSWPQVSLQD 20

RESULT 8
US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 51m & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-615-271-10

Query Match 100.0%; Score 112; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPQVSLQD 20
Db 1 IVNGEAVPGSWPQVSLQD 20

RESULT 9
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; REDUCED PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 51m & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:

```



```

; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-659-10

Query Match 100.0%; Score 112; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVFGSNPWQVSLQD 20
Db 1 INVGEAVFGSNPWQVSLQD 20

RESULT 10
US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,659
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155

```

```

; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-659-10

Query Match 100.0%; Score 112; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVFGSNPWQVSLQD 20
Db 1 INVGEAVFGSNPWQVSLQD 20

RESULT 11
US-09-106-468-10
; Sequence 10, Application US/09106468
; Patent No. 6114125
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,468
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REFERENCE/DOCKET NUMBER: 1038-825
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-106-468-10

Query Match 100.0%; Score 112; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 INVGEAVFGSNPWQVSLQD 20
Db 1 INVGEAVFGSNPWQVSLQD 20

```

```

RESULT 12
US-09-106-466A-10
; Sequence 10, Application US/09106466A
; Patent No. 6147057
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,466A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-826
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; FAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-106-466A-10
Query Match 100.0%; Score 112; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVNGEEAVPGSNPQVSLQD 20
Db 1 IVNGEEAVPGSNPQVSLQD 20

RESULT 13
US-09-106-467-10
; Sequence 10, Application US/09106467
; Patent No. 6153580
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; NUMBER OF SEQUENCES: 23
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: de Faire, Johan
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,540A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/600,273
; FILING DATE: 08-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen

```

```

; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,467
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; FAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-106-467-10
Query Match 100.0%; Score 112; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVNGEEAVPGSNPQVSLQD 20
Db 1 IVNGEEAVPGSNPQVSLQD 20

RESULT 14
US-08-385-540A-16
; Sequence 16, Application US/08385540A
; Patent No. 5945102
; GENERAL INFORMATION:
; APPLICANT: de Faire, Johan
; TITLE OF INVENTION: Wound Care With Multifunctional
; NUMBER OF SEQUENCES: 20
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,540A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/600,273
; FILING DATE: 08-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen

```

; REGISTRATION NUMBER: 29,135
 ; REFERENCE/DOCKET NUMBER: 314572-101A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 609-520-3214
 ; TELEFAX: 609-520-3259
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-385-540A-16

Query Match 97.3%; Score 109; DB 2; Length 20;
 Best Local Similarity 95.0%; Pred No. 1.9e-10;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEEAVFGSPWQVSLQD 20
 |||||:|||||
 Db 1 IVNGEDAVFGSPWQVSLQD 20

RESULT 15
 US-08-600-273A-16
 ; Sequence 16, Application US/08600273A
 ; Patent No. 5958406
 ; GENERAL INFORMATION:
 ; APPLICANT: de Faire, Johan
 ; APPLICANT: Franklin, Richard L.
 ; APPLICANT: Johnson, John
 ; TITLE OF INVENTION: Acne Treatment With Multifunctional
 ; TITLE OF INVENTION: Enzyme
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dechert Price & Rhoads
 ; STREET: 997 Lenox Drive, Building 3, Suite 210
 ; CITY: Lawrenceville
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 08543
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION NUMBER: US/08/600,273A
 ; APPLICATION NUMBER: US/08/600,273A
 ; FILING DATE: 08-FEB-1996
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/486,820
 ; FILING DATE: 07-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/385,540
 ; FILING DATE: 08-FEB-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bloom, Allen
 ; REGISTRATION NUMBER: 29,135
 ; REFERENCE/DOCKET NUMBER: 314572-101C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 609-620-3214
 ; TELEFAX: 609-620-3259
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-600-273A-16

Query Match 97.3%; Score 109; DB 2; Length 20;

Best Local Similarity 95.0%; Pred No. 1.9e-10;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 IVNGEEAVFGSPWQVSLQD 20
 |||||:|||||
 Db 1 IVNGEDAVFGSPWQVSLQD 20
 Search completed: February 12, 2003, 10:30:10
 Job time : 4.29851 secs

THIS PAGE BLANK (USPTO)


```

/ PRIOR APPLICATION NUMBER: PCT/US00/05989
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 928
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 643
/ LENGTH: 146
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (2)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (94)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (126)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (130)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (133)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (137)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (143)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-643

Query Match 97.3%; Score 109; DB 10; Length 146;
Best Local Similarity 95.0%; Pred. No. 1.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSWPMQVSLQD 20
Db 38 IVNGEDAVPGSWPMQVSLQD 57

RESULT 3
US-09-888-615-96
/ Sequence 96, Application US/09888615
/ Patent No. US2002064856A1
/ GENERAL INFORMATION:
/ APPLICANT: FLOWMAN, GREGORY
/ APPLICANT: WHITE, DAVID
/ APPLICANT: CAENEDEL, SEAN
/ APPLICANT: CHARVETZAK, GLEN
/ APPLICANT: MANNING, GREGARD
/ APPLICANT: SUDARSANAM, SUCHA
/ TITLE OF INVENTION: NOVEL PROTEASES
/ FILE REFERENCE: 038602/1214
/ CURRENT APPLICATION NUMBER: US/09/888,615
/ CURRENT FILING DATE: 2001-06-26
/ PRIOR APPLICATION NUMBER: 60/214,047
/ PRIOR FILING DATE: 2000-06-26
/ NUMBER OF SEQ ID NOS: 150
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 96
/ LENGTH: 263
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-888-615-96

Query Match 97.3%; Score 109; DB 10; Length 263;
Best Local Similarity 95.0%; Pred. No. 2.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSWPMQVSLQD 20

```

```

Db 34 IVNGEDAVPGSWPMQVSLQD 53

RESULT 4
US-09-925-297-529
/ Sequence 529, Application US/09925297
/ Patent No. US20020081659A1
/ GENERAL INFORMATION:
/ APPLICANT: ROSEN ET AL.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA105
/ CURRENT APPLICATION NUMBER: US/09/925,297
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05989
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 928
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 529
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-925-297-529

Query Match 86.6%; Score 97; DB 10; Length 192;
Best Local Similarity 90.0%; Pred. No. 9e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSWPMQVSLQD 20
Db 30 IVNGENAVLGSWPMQVSLQD 49

RESULT 5
US-09-923-779-152
/ Sequence 152, Application US/09923779
/ Patent No. US2002007672A1
/ GENERAL INFORMATION:
/ APPLICANT: FYLE, RICH A.
/ APPLICANT: KALOG, MICHAEL D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
/ FILE REFERENCE: 210121.553
/ CURRENT APPLICATION NUMBER: US/09/923,779
/ CURRENT FILING DATE: 2001-08-06
/ NUMBER OF SEQ ID NOS: 155
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 152
/ LENGTH: 270
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-923-779-152

Query Match 78.6%; Score 88; DB 10; Length 270;
Best Local Similarity 78.9%; Pred. No. 2.3e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGEEAVPGSWPMQVSLQ 19
Db 29 VVHGEDAVPYSWPMQVSLQ 47

RESULT 6
US-09-925-297-576
/ Sequence 576, Application US/09925297
/ Patent No. US20020081659A1
/ GENERAL INFORMATION:
/ APPLICANT: ROSEN ET AL.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA105

```

```

; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 576
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (213)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (220)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (214)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-297-576

Query Match 76.8%; Score 86; DB 10; Length 269;
Best Local Similarity 78.9%; Pred. No. 4.3e-05;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSNPMQVSLQ 19
: ||||| |||||
DB 34 VVGGEARPNISPMQVSLQ 52
: ||||| |||||

RESULT 7
US-09-925-297-695
; Sequence 695; Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-297-695

Query Match 76.8%; Score 86; DB 10; Length 273;
Best Local Similarity 78.9%; Pred. No. 4.4e-05;

```

```

Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSNPMQVSLQ 19
: ||||| |||||
DB 32 VVGGEDAVPSNPMQVSLQ 50
: ||||| |||||

RESULT 8
US-10-036-371-4
; Sequence 4; Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: EMMANARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; FILE REFERENCE: PCT/US99/048560
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Gadus sp.
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (11)
; OTHER INFORMATION: S or T
; NAME/KEY: MOD RES
; LOCATION: (13)
; OTHER INFORMATION: S, P or Y
; NAME/KEY: MOD RES
; LOCATION: (20)
; OTHER INFORMATION: D or Q
; US-10-036-371-4

Query Match 75.9%; Score 85; DB 12; Length 20;
Best Local Similarity 84.2%; Pred. No. 4.6e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVNGEEAVPGSNPMQVSLQ 19
: ||||| |||||
DB 1 IVNGEEAVPEKXWQVSLQ 19
: ||||| |||||

RESULT 9
US-09-988-975A-1
; Sequence 1; Application US/09988975A
; Patent No. US20020119531A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: PROSTATE-ASSOCIATED PROTEASE ANTIBODY
; FILE REFERENCE: PF-0227-2 CIP
; CURRENT APPLICATION NUMBER: US/09/988,975A
; CURRENT FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: site feature
; OTHER INFORMATION: Incyte ID No. US20020119531A1 556016
; NAME/KEY: unsure

```

```

; LOCATION: 235
; OTHER INFORMATION: unknown or other
US-09-986-978A-1

Query Match 75.0%; Score 84; DB 10; Length 283;
Best Local Similarity 77.8%; Pred. No. 8.6e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPMQVSL 18
Db 47 IVGESALPGAWPMQVSL 64

RESULT 10
US-09-981-353-23
; Sequence 23, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Laasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 991163CD1
US-09-981-353-23

Query Match 75.0%; Score 84; DB 9; Length 384;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPMQVSL 18
Db 148 IVGESALPGAWPMQVSL 165

RESULT 11
US-10-012-896-934
; Sequence 934, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-934

Query Match 75.0%; Score 84; DB 9; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPMQVSL 18
Db 157 IVGESALPGAWPMQVSL 174

RESULT 13
US-09-895-814-934

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-934

Query Match 75.0%; Score 84; DB 9; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPMQVSL 18
Db 157 IVGESALPGAWPMQVSL 174

RESULT 12
US-09-895-793-934
; Sequence 934, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-934

Query Match 75.0%; Score 84; DB 9; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVNGEAVPGSWPMQVSL 18
Db 157 IVGESALPGAWPMQVSL 174

RESULT 13
US-09-895-814-934

```



```

; Sequence 934, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carola
; APPLICANT: Fanger, Gary R.
; APPLICANT: Fetter, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-934

```

```

Query Match          75.0%; Score 84; DB 9; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVNGEAVPGSWPQVSL 18
    |||||:|||||
Db 157 IVGESALPGMPQVSL 174

```

```

RESULT 14
US-09-759-143-934
; Sequence 934, Application US/09759143
; Publication No. US200202248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/759,143

```

```

; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-934

```

```

Query Match          75.0%; Score 84; DB 10; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVNGEAVPGSWPQVSL 18
    |||||:|||||
Db 157 IVGESALPGMPQVSL 174

```

```

RESULT 15
US-09-780-669-934
; Sequence 934, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 10121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 934
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-934

```

```

Query Match          75.0%; Score 84; DB 10; Length 393;
Best Local Similarity 77.8%; Pred. No. 0.00012;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVNGEAVPGSWPQVSL 18
    |||||:|||||
Db 157 IVGESALPGMPQVSL 174

```

```

Search completed: February 12, 2003, 10:31:06
Job time : 3.04478 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:03:55, Search time 2.68657 Seconds
(without alignment)
198.395 Million cell updates/sec

Title: US-10-036-371-7

Perfect score: 21

Sequence: 1 AAPF 4

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	5 AAP40459	N-terminal polyhyd
2	21	100.0	4	11 AAR04012	Peptidase substrat
3	21	100.0	4	12 AAR10887	Peptide component
4	21	100.0	4	12 AAR11223	Ketone analogue pr
5	21	100.0	4	13 AAR29964	Chetapsin G inhibi
6	21	100.0	4	14 AAR34258	Chromogenic peptid
7	21	100.0	4	14 AAR38410	Cathepsin G inhibi
8	21	100.0	4	14 AAR44109	Aminonaphthalene p
9	21	100.0	4	15 AAR52024	Bacillus alkali pr
10	21	100.0	4	15 AAR62223	Serine protease in

11	21	100.0	4	15 AAS53781	Sequence of intern
12	21	100.0	4	16 AAR72920	Substrate for pept
13	21	100.0	4	16 AAR72876	Substrate for pept
14	21	100.0	4	16 AAR77315	Porphyromonas ging
15	21	100.0	4	16 AAR77196	Cell proliferation
16	21	100.0	4	17 AAR85709	Degradable peptide
17	21	100.0	4	17 AAR52606	Serine protease-in
18	21	100.0	4	18 AAR30747	Substrate for seri
19	21	100.0	4	18 AAR24567	Proteinase site of
20	21	100.0	4	18 AAR06165	Synthetic substrat
21	21	100.0	4	18 AAR12810	B. subtilis subtitl
22	21	100.0	4	19 AAR76994	B. subtilis subtitl
23	21	100.0	4	19 AAR76994	Synthetic oligopep
24	21	100.0	4	19 AAR61327	Indicator for dete
25	21	100.0	4	19 AAR51610	Alkaline protease
26	21	100.0	4	19 AAR51610	Alkaline protease
27	21	100.0	4	20 AAR48424	Peptide comprising
28	21	100.0	4	20 AAR07072	Chymotrypsin enzym
29	21	100.0	4	20 AAR84189	Substrate peptide
30	21	100.0	4	21 AAR20774	Synthetic PPIase s
31	21	100.0	4	21 AAR03092	Model substrate pe
32	21	100.0	4	21 AAR01922	Protease peptide s
33	21	100.0	4	21 AAY91025	Chymase inhibitor
34	21	100.0	4	21 AAY97813	Chymotrypsin subst
35	21	100.0	4	21 AAY80465	Subtilisin substra
36	21	100.0	4	21 AAY76808	Cathepsin G peptid
37	21	100.0	4	21 AAY78789	Peptidylprolyl cis
38	21	100.0	4	21 AAY49452	Peptide substrate u
39	21	100.0	4	21 AAY59632	Peptide substrate
40	21	100.0	4	21 AAY00792	Model peptide subs
41	21	100.0	4	22 AAG64792	Model peptide subs
42	21	100.0	4	22 AAG64496	Human peptide #2
43	21	100.0	4	22 AAG64496	Human peptide #2
44	21	100.0	4	22 AAS04342	Peptide #2. Unide
45	21	100.0	4	22 AAS98623	

ALIGNMENTS

RESULT 1
AAP40459
ID AAP40459 standard; Protein; 4 AA.
XX
XX AAP40459;
DT 27-NOV-1991 (first entry)
XX
DE N-terminal polyhydroxyalkanoyl peptide.
KW Polyhydroxyalkanoyl peptide; protease; peptidase; prodrug.
XX
XX EP126685-A.
XX
PD 28-NOV-1985.
XX
PF 15-MAY-1984; 84EP-0400984.
XX
PR 16-MAY-1983; 83FR-0308051.
XX
XX (CNRS) CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE.
PI
PI Monsigny M, Mayer R;
XX
XX WP1; 1984-256065/48.
XX
PT N-terminal polyhydroxyalkanoyl peptide and aminoacid derivs. - opt.
PT having C-terminal active gpe., e.g. chloroquine, are water-soluble
PT antitumour or antiparasitic cpds. and protease targets.
XX
PS Claim 11; page 15; 20pp; french.
XX
XX The peptide is polyhydroxyalkanoyl-L-Ala-L-Ala-L-Pro-L-Phe-X. X is

CC R2-NH- derived from an aromatic amine, eg P-nitro-aniline, alpha-
 CC or beta-naphthylamine, or it is RNH derived from an amino drug, eg
 CC daunorubicin or chloroquine, or it is a gp. which may confer
 CC activity, eh H, OH, CH₂Cl. When X is the residue of a drug, the
 CC cpd. is a prodrug in which the peptide is the substrate for a
 CC specific protease secreted by the target cell for the active drug,
 CC eg tumour cells or microorganism pathogens, eg Plasmodium
 CC falciparum. Due to the polyhydroxy gp., the cpds. are very soluble
 CC in water or buffer soln. and are able to be used at high concn. The
 CC cpds. also allow the detection of proteases and peptidases and
 CC allow easy determination of the best substrate for a particular
 CC protease.
 CC Sequence 4 AA;

Query Match 100.0%; Score 21; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 2
 AAR04012
 ID AAR04012 standard; peptide; 4 AA.

XX AAR04012;

XX 19-NOV-1991 (first entry)

XX Peptidase substrate analogues having peptidase inhibition activity.

XX Elastase; plasminogen activator; cathepsin; calpain; enkephalinase;
 KW Angiotensin converting enzyme.

XX Synthetic.

PH Key Location/Qualifiers
 FT Modified-site /Label: Terminal modified from -COOH to -COCOR
 FT /note= "R = alkyl group or -H"

XX EP363284-A.

XX 11-APR-1990.

XX 06-OCT-1989; 89EP-0402762.

XX 07-OCT-1988; 88US-0254762.

XX 06-OCT-1989; 89EP-0402763.

XX (RICH) MERRELL DOW PHARM INC.

XX Bey P, Angelastro M, Mehdi S;

XX WPI; 1990-109579/15.

XX New peptidase substrate analogue cpds. - useful as protease
 PT inhibitors in treatment of disease states.

XX Claim 5; Page 26; 33pp; English.

XX The analogues may be useful in treatment of a variety of disease
 CC states. The scissile amide group is replaced with H or a substituted
 CC carbon moiety effectively inhibiting the activity of peptidases such
 CC as elastase, plasmin thrombin, urokinase etc.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 11; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.9e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 3
 AAR10887
 ID AAR10887 standard; Protein; 4 AA.

XX AAR10887;

XX 10-APR-1991 (first entry)

XX Peptide component of pentafluoroethylcarbonyl analogue.

XX Protease inhibitor; rheumatoid arthritis; thrombosis; psoriasis;
 KW male contraceptive.

XX Synthetic.

XX EP410411-A.

XX 30-JAN-1991.

XX 25-JUL-1990; 90EP-0114250.

XX 26-JUL-1989; 89US-0385624.

XX (RICH) MERRELL DOW PHARM INC.

XX Bey P, Peet NP, Angelastro MR, Mehdi S;

XX WPI; 1991-030811/05.

XX Novel serine- carboxylic acid-and metallo-proteinase-inhibitors
 PT (broad range of proteinase(s) in treating rheumatoid
 PT arthritis, thrombosis and psoriasis, also is a male contraceptive

XX Claim 5; page 25; 40pp; English.

XX This peptide is the R1 gp. of the cpd. of formula: R1NHCH(R2)-
 CC COCF₂CF₃. This cpd. is a peptidase substrate analogue and is a
 CC specific enzyme inhibitor for a range of proteases, e.g. serine-,
 CC carboxylic acid-, and metallo-proteases. It is useful in the
 CC treatment of rheumatoid arthritis, thrombosis and psoriasis and is
 CC also used as a male contraceptive. See also AAR10876-83, AAR10886 and
 CC AAR10888.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 12; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAPF 4
 ||||
 Db 1 AAPF 4

RESULT 4
 AAR11223
 ID AAR11223 standard; Protein; 4 AA.

XX AAR11223;

XX 24-MAY-1991 (first entry)

XX Ketone analogue protease inhibitor #2.

XX protease inhibitor; antiinflammatory agent; hypotensive; analgesic;
 KW antiproliferative agent; anticidemyelinating agent; antithrombotic.

```

OS XX      Synthetic.
FH XX      Key      Location/Qualifiers
FT XX      Modified-site 1
FT XX      /label= Methoxysuccinyl-Ala
FT XX      Modified-site 4
FT XX      /label= Phe-carboxyoxalyl
XX XX
PN XX      EP417721-A.
XX XX
PD XX      20-MAR-1991.
XX XX
PF XX      11-SEP-1990; 90EP-0117461.
XX XX
PR XX      11-SEP-1989; 89US-0405491.
XX XX
PA XX      (RICH ) MERRELL DOW PHARM INC.
XX XX
PI XX      Flynn GA, Bey P;
XX XX      WPI; 1991-081980/12.
XX XX
XX XX      New ketone analogue peptidase and isomerase inhibitors - for
XX XX      inhibition of leukocyte elastase, cathepsin G, thrombin,
XX XX      chymotrypsin, plasmin etc.
XX XX
PS XX      Claim 5; Page 26; 50pp; English.
XX XX
CC XX      This peptide is a specific example of a highly generic protease
CC XX      inhibitor useful for medical purposes. The peptide analogues include
CC XX      inhibitors of urokinase, renin, cathepsin D etc. which can be used
CC XX      as anti-proliferative agents and abortifacients, hypotensives,
CC XX      anti-inflammatory and antidemylinating agents, respectively.
CC XX      See also AAR11222 and AAR11224-R11218.
XX XX
SQ XX      Sequence 4 AA;

Query Match      100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 1 AAPF 4

RESULT 5
AAR29964
ID AAR29964 standard; peptide; 4 AA.
XX
AC AAR29964;
XX
DT 19-APR-1993 (first entry)
XX
DE Cathepsin G inhibiting fragment.
XX
XX Cathepsin G; elastase; connective tissue; degradation; protease;
XX Gout; rheumatoid arthritis; emphysema; AKUS;
XX adult respiratory distress syndrome; para-phenylene.
XX
OS Synthetic.
XX
FH XX      Key      Location/Qualifiers
FT XX      Modified-site 1
FT XX      /note= "N-terminal amino acids of the peptides of
FT XX      AAR29963-64 are pref. linked by
FT XX      -C(O)-phenylene-C(O)-, esp. wherein the
FT XX      phenylene is a para-phenylene gp."
FT XX      Modified-site 4
FT XX      /note= "C-terminal Phe is in keto form. i.e. OH
FT XX      replaced by CO1W, pref. CF3 or CF2CF3"
XX XX      WO22020357-A.

XX XX      26-NOV-1992.
XX XX
XX XX      21-APR-1992; 92WO-US03288.
XX XX
XX XX      23-MAY-1991; 91US-0704449.
XX XX
XX XX      (RICH ) MERRELL DOW PHARM INC.
XX XX
XX XX      Angelastro MR, Bey P, Doherty NS, Janusz MJ, Mehdi S;
XX XX      Peet NP;
XX XX      WPI; 1992-415461/50.
XX XX
XX XX      New peptide derivs. used as cathepsin G and elastase inhibitors -
XX XX      for treating gout, rheumatoid arthritis, inflammatory disorders,
XX XX      emphysema and adult respiratory distress syndrome
XX XX
XX XX      Claim 13-17; Page 52; 53pp; English.
XX XX
XX XX      This sequence is an example of a highly generic formula.
XX XX      Inhibitors of cathepsin G and elastase for preventing connective
XX XX      tissue degradation are chemically linked inhibitors of the proteases
XX XX      elastase (pref. the peptide of AAR29963, or Lys(2CBz)-Pro-Val or
XX XX      Val-Pro-Val) and cathepsin G (pref. the peptide of AAR29964, or
XX XX      Val-Pro-Phe or Phe).
XX XX      The N-terminal amino acids of the peptides of AAR29963-64 are pref.
XX XX      linked by -C(O)-Phenylene-C(O)-, esp. wherein the phenylene is a
XX XX      para-phenylene gp.
XX XX      The cpds. have an anti-inflammatory effect useful in the treatment
XX XX      of gout, rheumatoid arthritis and other inflammatory diseases and
XX XX      to enhance the anti-inflammatory tissue damage. They can also be used
XX XX      in the treatment of emphysema and adult respiratory distress syndrome.
XX XX
SQ XX      Sequence 4 AA;

Query Match      100.0%; Score 21; DB 13; Length 4;
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 1 AAPF 4

RESULT 6
AAR34258
ID AAR34258 standard; peptide; 4 AA.
XX
AC AAR34258;
XX
DT 19-AUG-1993 (first entry)
XX
DE Chromogenic peptide substrate for peptidase assay.
XX
XX Alkaline protease; detergent; stability.
XX
OS Synthetic.
XX
FH XX      Key      Location/Qualifiers
FT XX      Modified-site 1
FT XX      /note= "N-succinyl alanine"
FT XX      Modified-site 4
FT XX      /note= "Phe-p-nitroanilide"
XX XX      WO9307276-A.
XX XX
PD XX      15-APR-1993.
XX XX
XX XX      07-OCT-1992; 92WO-US08341.
XX XX
XX XX      08-OCT-1991; 91US-0772087.
XX XX

```

4
.
.
2

CC followed by reaction with 5-aminonaphthalene-1-(N-propyl)sulphamide.
 CC The new compounds are used as detector groups in quantitative
 CC determination of chymotrypsin.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 14; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 9

AAR52024
 ID AAR52024 standard; peptide; 4 AA.

XX AC AAR52024;

XX DT 28-NOV-1994 (first entry)

XX DE Bacillus alkali protease substrate.

XX KW Bacillus NKS-21 alkali protease; hydrolysis; thermostable enzyme;
 XX detergent stable.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1

FT /label= OTHER

FT /note= "Suc-Ala"

FT Modified-site 1

FT /label= OTHER

FT /note= "Phe-MCA"

XX PN JP06070765-A.

XX PD 15-MAR-1994.

XX PF 08-OCT-1992; 92JP-0296360.

XX PR 10-JUL-1992; 92JP-0207302.

XX PA (SHOW) SHOWA DENKO KK.

XX DR WPI; 1994-128672/16.

XX FT New alkali protease stable to heat and detergent - useful as
 XX industrial enzyme, eg in washing compositions

XX PS Claim 1; Page 2, 10pp; Japanese.

XX CC A novel alkali protease obtained from Bacillus NKS-21 is defined by
 CC its physico-chemical properties including substrate specificity for
 CC the tetrapeptide AAR52024.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 10

AAR46223
 ID AAR46223 standard; peptide; 4 AA.

XX AAR46223;
 AC 04-AUG-1994 (first entry)
 DT Serine protease inhibitor tetrapeptide.
 DE Prevention; schistosomiasis; parasite; infection; prevention;
 KW parasitic penetration; skin; cercariae; anti-penetrant.
 XX Synthetic.

XX OS Key Location/Qualifiers

XX FH Modified-site 1

FT /note= "BG(peptide blocking gp.) attached"

FT Modified-site 4

FT /note= "PI(protease inhibitor), other than

FT Chloromethyl Ketone, attached"

XX PN US5284829-A.

XX PD 08-FEB-1994.

XX PF 26-NOV-1991; 91US-0798565.

XX PR 26-NOV-1991; 91US-0798565.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Cohen FE, McKerrow JH;

XX DR WPI; 1994-056364/07.

XX PT Synthetic tetra-peptide(s) with an N-terminal blocking gp. and
 FT C-terminal enzyme inhibitor - can be formulated into soaps and
 FT sprays and used to prevent schistosomal skin penetration

XX PS Disclosure; Page 7; 35pp; English.

XX CC The sequence is that of a synthetic tetrapeptide serine protease
 CC inhibitor which can be used to prevent schistosome parasite
 CC infection. It may be used in a formulation as a soap, lotion,
 CC cream, spray, etc. to stop parasitic penetration of the skin.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 11

AAR53781
 ID AAR53781 standard; Peptide; 4 AA.

XX AC AAR53781;

XX DT 29-DEC-1994 (first entry)

XX DE Sequence of internal fragment of a chromogenic substrate for
 DE cathepsin G.

XX KW Enzyme; cathepsin G; protease; chromogenic substrate.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /label= Suc-Ala

```

FT Modified-site 4 /note= "Succinyl group"
FT FT /label= Phe-pNA
FT FT /note= "p-nitroanil"
XX PN W09412637-A.
XX XX
XX PD 09-JUN-1994.
XX XX
XX PF 01-DEC-1993; 93WO-US11696.
XX XX
XX PR 02-DEC-1992; 92US-0985692.
XX PR 19-NOV-1993; 93US-0155331.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX XX
XX PI Foster DC, Norris K, Sprecher CA;
XX DR
XX DR WPI; 1994-200265/24.
XX XX
XX PR New human Kunitz type protease inhibitor and related DNA - for
XX FT treating pancreatic and other disorders involving serine
XX FT protease, also new amyloid protein precursor homologues including
XX FT the inhibitor in its sequence
XX XX
XX PS Example; Page 63; 70pp; English.
XX XX
XX CC Protease inhibitory profiles of the Kunitz inhibitors
XX CC were determined for a variety of proteases using
XX CC a variety of chromogenic substrates and compared to the inhibitory
XX CC activity shown by the Kunitz-type inhibitor domain of the amyloid
XX CC protein precursor and bovine aprotinin. The substrate AAR53780
XX CC was tested using the protease leukocyte elastase.
XX XX
XX SQ Sequence 4 AA;
XX
XX Query Match 100.0%; Score 21; DB 15; Length 4;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy 1 AAPP 4
Db 1 AAPP 4
XX
XX RESULT 12
XX AAR72920
XX ID AAR72920 standard; Peptide; 4 AA.
XX AC AAR72920;
XX XX
XX DT 29-NOV-1995 (first entry)
XX DE
XX DE Substrate for peptidyl prolyl cis trans isomerase alpha.
XX XX
XX KW Escherichia coli; protein conformation; folding; acceleration;
XX KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
XX KW catalysis; isomerisation; prolyl peptide bond.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-succinyl-Ala"
XX FT Modified-site 4 /note= "Phe-MCA (sic)"
XX FT
XX PN EP647714-A.
XX XX
XX PD 12-APR-1995.
XX PF
XX PF 19-JUL-1990; 90EP-0307914.
XX XX
XX PR 19-JUL-1989; 89JP-0184738.
XX PR 06-OCT-1989; 89JP-0260244.
XX PR 29-DEC-1989; 89JP-0344705.
XX XX
XX PA (TOFU ) TONEN CORP.
XX XX
XX PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX DR
XX WPI; 1995-140755/19.

```

```

XX 19-JUL-1989; 89JP-0184738.
XX PR 06-OCT-1989; 89JP-0260244.
XX PR 29-DEC-1989; 89JP-0344705.
XX XX
XX PA (TOFU ) TONEN CORP.
XX XX
XX PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX DR
XX DR WPI; 1995-140756/19.
XX XX
XX PR New E.coli peptidyl prolyl cis trans isomerase beta - used to
XX FT accelerate the folding of proteins, partic. for activation of
XX FT inactive recombinant proteins
XX XX
XX PS Disclosure; Page 5; 85pp; English.
XX XX
XX CC To measure the activity of E. coli PPIase-alpha (peptidyl prolyl cis
XX CC trans isomerase alpha), the PPIase and Csa (cyclosporine A) are
XX CC added in the cell and mixed for 1 minute. Then, this peptide is
XX CC added and incubated, and chymotrypsin added to start the reaction.
XX CC The effect of Csa on the inhibition of the PPIase activity can be
XX CC detected by varying the amt. of the Csa added. The inventors are
XX CC claiming a PPIase-beta.
XX XX
XX SQ Sequence 4 AA;
XX
XX Query Match 100.0%; Score 21; DB 16; Length 4;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy 1 AAPP 4
Db 1 AAPP 4
XX
XX RESULT 13
XX AAR72876
XX ID AAR72876 standard; Peptide; 4 AA.
XX AC AAR72876;
XX XX
XX DT 29-NOV-1995 (first entry)
XX DE
XX DE Substrate for peptidyl prolyl cis trans isomerase alpha.
XX XX
XX KW Escherichia coli; protein conformation; folding; acceleration;
XX KW PPIase-alpha; peptidyl prolyl cis trans isomerase alpha;
XX KW catalysis; isomerisation; prolyl peptide bond.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-succinyl-Ala"
XX FT Modified-site 4 /note= "Phe-MCA (sic)"
XX FT
XX PN EP647713-A.
XX XX
XX PD 12-APR-1995.
XX PF
XX PF 19-JUL-1990; 90EP-0307914.
XX XX
XX PR 19-JUL-1989; 89JP-0184738.
XX PR 06-OCT-1989; 89JP-0260244.
XX PR 29-DEC-1989; 89JP-0344705.
XX XX
XX PA (TOFU ) TONEN CORP.
XX XX
XX PI Hayano T, Katou S, Maki N, Suzuki M, Takahashi N;
XX DR
XX WPI; 1995-140755/19.

```



```

XX New E.coli peptidyl prolyl cis trans isomerase alpha - used to
PT accelerate the folding of proteins, partic. for activation of
PT inactive recombinant proteins
XX
XX Disclosure; Page 5; 85pp; English.
XX
CC To measure the activity of E. coli PPIase-alpha (peptidyl prolyl cis
CC trans isomerase alpha), the PPIase and Csa (cyclosporine A) are
CC added in the cell and mixed for 1 minute. Then, this peptide is
CC added and incubated, and chymotrypsin added to start the reaction.
CC The effect of Csa on the inhibition of the PPIase activity can be
CC detected by varying the ant. of the Csa added. The inventors are
CC claiming the PPIase-alpha.
XX
XX Sequence 4 AA;
SQ
Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAPF 4
XX ||||
XX Db 1 AAPF 4
XX
RESULT 15
AAR77196
ID AAR77196 standard; peptide; 4 AA.
XX
XX AAR77196;
AC
XX 27-FEB-1996 (first entry)
DI
XX Cell proliferation enzyme proteinase activity substrate peptide #1.
XX
XX Proteinase; cell growth-stimulating protein; hydrolysis;
KW macrophage chemotactic action; serine protease inhibitor; wound;
KW gastric ulcer; leg ulcer; bed sore.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /label= Succinyl-Ala
FT Modified-site 4
FT Modified-site 4 /note= "Modified by 4-methyl-coumaryl-7-amide"
XX
XX EP661293-A2.
XX
XX 05-JUL-1995.
XX
XX 22-DEC-1994; 94BP-0120406.
XX
XX 29-DEC-1993; 93JP-0351225.
XX
XX (SANW ) SANWA KAGAKU KENKYUSHO CO.
XX
XX Awaya J, Uesaka H, Watanabe K, Yamaguchi T;
XX WPI; 1995-233274/31.
XX
XX Protein derived from Clostridium perfringens FERM BP-4584 -
XX stimulates cell proliferation and has macrophage chemotactic action
XX
XX Example; Page 8; 14pp; English.
XX
XX The sequences given in AAR77196-200 are peptides which were used to
XX demonstrate the proteinase activity of the cell growth-stimulating
XX protein of the invention. The protein was seen to selectively
XX hydrolyse synthetic substrates having an aromatic amino acid at the
XX C-terminal. The response to the peptide given in AAR77196 was
XX particularly high. The cell growth-stimulating protein has a mol.
XX wt. of 420 +/- 40 kD and a single subunit mol. wt. of 130 +/- 20 kD.
XX It has an isoelectric point of 4.8 and has cell growth stimulating
XX action and macrophage chemotactic action, as well as proteinase
XX activity. Its enzymatic activity decreases in the presence of a
XX serine protease inhibitor, increase in the presence of various metal
XX ions, and is stabilised in the presence of calcium ion. The protein is
XX particularly useful for the treatment of wounds, gastric and leg ulcers,
XX eg. bed sores.
XX
XX Sequence 4 AA;
SQ
Query Match 100.0%; Score 21; DB 16; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAPF 4
XX ||||
XX Db 1 AAPF 4
XX
RESULT 14
AAR77315
ID AAR77315 standard; Protein; 4 AA.
XX
XX AAR77315;
AC
XX 28-FEB-1996 (first entry)
DI
XX Porphyromonas gingivalis protease substrate.
XX
XX Protease; periodontal disease; pathogenic microbe, diagnosis;
KW substrate.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note= "Suc-Ala"
FT Modified-site 4 /note= "MCA-Phe"
XX
XX JP07135973-A.
XX
XX 30-MAY-1995.
XX
XX 15-NOV-1993; 93JP-0307084.
XX
XX 15-NOV-1993; 93JP-0307084.
XX
XX (SUNR ) SUNTORY LTD.
XX
XX WPI; 1995-227397/30.
XX
XX An enzyme originated from a periodontal disease pathogenic microbe -
XX and an antibody against the enzyme, for the determination of the
XX progress and activity of the disease
XX
XX Example; Page 7; 15pp; Japanese.
XX
XX AAR77315 is a target substrate for a Porphyromonas gingivalis (a
XX periodontal disease pathogenic microbe) protease. An antibody
XX raised against the enzyme can be used to diagnose the
XX and progress of a periodontal disease, caused by a pathogenic
XX microbe.
XX
XX Sequence 4 AA;
SQ

```

Wed Feb 12 11:59:44 2003

us-10-036-371-7.rag

Page 8

DB 1 APP 4

Search completed: February 12, 2003, 10:22:28
Job time : 4.68657 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 1.00299 Seconds
(without alignments)
383.399 Million cell updates/sec

Title: US-10-036-371-7

Perfect score: 21

Sequence: 1 AAPF 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	5	2 JS0319	subesophageal gang
2	21	100.0	42	2 843963	hypothetical prote
3	21	100.0	47	2 PN0607	cytochrome-c oxida
4	21	100.0	50	2 AD3575	hypothetical prote
5	21	100.0	63	1 OSB08A	cytochrome-c oxida
6	21	100.0	63	2 S10303	cytochrome-c oxida
7	21	100.0	75	2 AB2778	hypothetical prote
8	21	100.0	79	2 C69125	hypothetical prote
9	21	100.0	80	2 C36953	probable metal-bin
10	21	100.0	86	2 E90905	hypothetical prote
11	21	100.0	88	2 C62429	flg10 protein (AF30
12	21	100.0	88	2 AC2429	flagellar biosynth
13	21	100.0	90	2 C02854	male accessory gla
14	21	100.0	90	2 S30428	Mat26Ab protein -
15	21	100.0	90	2 S30412	male accessory gla
16	21	100.0	90	2 C30430	male accessory gla
17	21	100.0	94	2 C98239	hypothetical prote
18	21	100.0	97	2 AB1502	probable cell surf
19	21	100.0	100	2 G86174	hypothetical prote
20	21	100.0	101	2 D90761	PE-family protein
21	21	100.0	102	2 B86976	probable PE protei
22	21	100.0	102	2 H70898	hypothetical prote
23	21	100.0	107	2 F75397	hypothetical prote
24	21	100.0	108	2 F72507	protein t17H7.7 [i
25	21	100.0	109	2 G86433	probable transcrip
26	21	100.0	112	2 H96031	anti-SS-A/Ro 60K p
27	21	100.0	122	2 FC4279	anti-SS-A/Ro 60K p
28	21	100.0	122	2 FC4280	hypothetical prote
29	21	100.0	122	2 C63501	hypothetical prote

ALIGNMENTS

RESULT 1

JS0319

subesophageal ganglion pentapeptide - house cricket

C:Species: Acheta domestica (house cricket)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000

C:Accession: JS0319

C:Wicker, C.; Wicker, C.

Comp. Biochem. Physiol. C 88, 185-187, 1987

A:Title: Isolation and structure of a peptide isolated from the subesophageal ganglion (

A:Reference: PMID:8190622

A:Accession: JS0319

A:Molecule type: protein

A:Residues: 1-5 <MIC>

Query Match 100.0%; Score 21; DA 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

DB 2 AAPF 5

RESULT 2

S43961

hypothetical protein (clone PR57) - Rhizobium sp. (strain NGR234) (fragment)

C:Species: Rhizobium

A:Variety: strain NGR234

C:Date: 20-Oct-1994 #sequence_revision 23-Feb-1996 #text_change 21-Aug-1998

C:Accession: S43963

R;Perret, X.; Pellay, R.; Bjourson, A.J.; Cooper, J.E.; Brenner, S.; Broughton, W.J.

Nucleic Acids Res. 22, 1335-1341, 1994

A:Title: Subtraction hybridisation and shot-gun sequencing: a new approach to identify s

A:Reference number: S43961; MUID:94248027; PMID:8190622

A:Accession: S43963

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-42 <PER>

A:Experimental source: strain NGR234

C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 100.0%; Score 21; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

DB 10 AAPF 13

RESULT 3

PN0607

cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - pig (fragment)
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C/Accession: PN0607
 E/Sillard, R.J.; Joenvall, H.; Mutt, V.
 Biochem. Biophys. Res. Commun. 195, 746-750, 1993
 A/Title: Characterization of porcine intestinal cytochrome c oxidase subunit VIIC, purified from pig
 A/Reference number: PN0607; MUID:93384597; PMID:8396926
 A/Accession: PN0607
 A/Molecule type: protein
 A/Residues: 1-47 <SL>
 A/Experimental source: intestine
 C/Superfamily: cytochrome-c oxidase chain VIIC
 C/Keywords: mitochondrion; oxidoreductase; respiratory chain

Query Match 100.0%; Score 21; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPP 4
 DB 34 AAPP 37

RESULT 4
 AD3575
 hypothetical protein BMEH10525 [imported] - Brucella melitensis (strain 16M)
 C/Species: Brucella melitensis
 C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C/Accession: AD3575
 R/DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Ios, T.; Ivanova, M.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesce, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A/Reference number: AD3252; PMID:1175668
 A/Accession: AD3575
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-50 <NN>
 A/Cross-references: GS:AE008918; PIDN:ALM53767.1; PID:gl7984695; GSPDB:GN00191
 A/Experimental source: strain 16M
 C/Genetics: BMEH10525
 A/Map position: 11

Query Match 100.0%; Score 21; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPP 4
 DB 44 AAPP 47

RESULT 5
 OSB08A
 cytochrome-c oxidase (EC 1.9.3.1) chain VIIC precursor [validated] - bovine
 N/Alternate names: cytochrome-c oxidase chain VIIa
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 31-May-1979 #sequence_revision 31-Dec-1992 #text_change 15-Sep-2000
 C/Accession: JH0473; S06597; S00498; H29968; S18834
 R/Aqua, M.S.; Bachman, N.J.; Lomax, M.I.; Grossman, L.I.
 Gene 104, 211-217, 1991
 A/Title: Characterization and expression of a cDNA specifying subunit VIIC of bovine cytochrome c oxidase
 A/Reference number: JH0473; MUID:92009215; PMID:1655579
 A/Accession: JH0473
 A/Molecule type: DNA
 A/Residues: 1-63 <AQUL>
 A/Cross-references: GS:X58823
 R/Aqua, M.S.; Lomax, M.I.; Schon, E.A.; Grossman, L.I.
 Nucleic Acids Res. 17, 8376, 1989
 A/Title: Nucleotide sequence of a cDNA for bovine cytochrome c oxidase subunit VIIC.
 A/Reference number: S06597; MUID:90045968; PMID:2554257

A/Accession: S06597
 A/Molecule type: mRNA
 A/Residues: 1-63 <AQUL>
 A/Cross-references: EMBL:X15725; NID:9277; PIDN:CAB57793.1; PID:g6015488
 R/Buse, G.; Steffens, G.J.
 Hoppe-Seyler's Z. Physiol. Chem. 359, 1005-1009, 1978
 A/Title: Studies on cytochrome c oxidase, II. The chemical constitution of a short polypeptide
 A/Reference number: A00498; MUID:79046803; PMID:213363
 A/Accession: A00498
 A/Molecule type: protein
 A/Residues: 17-63 <BUS>
 A/Experimental source: heart
 R/Yanamura, W.; Zhang, Y.Z.; Takamiya, S.; Capaldi, R.A.
 Biochemistry 27, 4909-4914, 1988
 A/Title: Tissue-specific differences between heart and liver cytochrome c oxidase.
 A/Reference number: A90531; MUID:89000697; PMID:2844245
 A/Accession: H29968
 A/Molecule type: protein
 A/Residues: 17-42 <YAN>
 R/Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-ito, H.; submitted to the Brookhaven Protein Data Bank, April 1996
 A/Reference number: A67451; PDB:1OCC
 A/Contents: annotation; X-ray crystallography, 2.8 angstroms, residues 17-63
 R/Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-ito, K.; Science 272, 1136-1144, 1996
 A/Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 angstroms
 A/Reference number: A57982; MUID:96216288; PMID:8638158
 A/Contents: annotation; X-ray crystallography, 2.8 angstroms
 C/Genetics: BMEH10525
 A/Cross-references: GS:AE008918; PIDN:ALM53767.1; PID:gl7984695; GSPDB:GN00191
 A/Experimental source: strain 16M
 C/Genetics: BMEH10525
 A/Map position: 11

Query Match 100.0%; Score 21; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPP 4
 DB 50 AAPP 53

RESULT 6
 S10303
 cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 11-Jun-1999
 C/Accession: S10303
 R/Akamat, M.; Grossman, L.I.
 Nucleic Acids Res. 18, 3645, 1990
 A/Title: Nucleotide sequence of a cDNA for mouse cytochrome c oxidase subunit VIIC.
 A/Reference number: S10303; MUID:93031494; PMID:2163523
 A/Accession: S10303
 A/Molecule type: mRNA
 A/Residues: 1-63 <AKA>
 A/Cross-references: EMBL:X52940; NID:g50524; PIDN:CAA37115.1; PID:g50525
 R/Akamat, M.; Grossman, L.I.
 Nucleic Acids Res. 17, 8376, 1989
 A/Title: Nucleotide sequence of a cDNA for bovine cytochrome c oxidase subunit VIIC.
 A/Reference number: S06597; MUID:90045968; PMID:2554257

C;Keywords: mitochondrion; oxidoreductase

```

Query Match      100.0%; Score 21; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFP 4
    ||||
Db 50 AAPF 53

RESULT 7
AB2778
hypothetical protein Atu1639 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AB2778
R;Wood, D.W.; Sekubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
stet, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB2778
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <KOR>
A;Cross-references: GB:AE008688; PIDN:AAU42640.1; PID:g17740071; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1639
A;Map position: circular chromosome

Query Match      100.0%; Score 21; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFP 4
    ||||
Db 50 AAPF 53

RESULT 8
CG9125
hypothetical protein MTH206 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: CG9125
R;Smith, D.R.; Doucette-Stamm, L.A.; Delouhery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Olu, D.; Spadafora, R.; Vicaria, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Iwami, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H. funct
A;Reference number: AE9000; MUID:98037514; PMID:9371463
A;Accession: CG9125
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-79 <MTH>
A;Cross-references: GB:AE000807; GB:AE000666; NID:g2621239; PIDN:AA84712.1; PID:g262125
C;Genetics:
A;Gene: MTH206
A;Start Codon: TTG

Query Match      100.0%; Score 21; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAFP 4
    ||||
Db 36 AAPF 39

```

```

RESULT 9
G96953
probable metal-binding protein [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: G96953
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, O.; Gibson, R.; Lee
J.; Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G96953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <KTR>
A;Cross-references: GB:AE001437; PIDN:AAK78418.1; PID:9J5023293; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0438

Query Match      100.0%; Score 21; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 63 AAPF 66

RESULT 10
E90905
hypothetical protein Ecs2213 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: E90905
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genor
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E90905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-86 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA935636.1; PID:g13361679; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain R1MD 050952
C;Genetics:
A;Gene: Ecs2213

Query Match      100.0%; Score 21; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
    ||||
Db 26 AAPF 29

RESULT 11
F97429
ILQ protein (AF300968) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C;Accession: F97429
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
Sci. 294, 2321-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: F97429
A;Status: preliminary

```

A:Molecule type: DNA
 A:Residues: 1-88 <KUR>
 A:CROSS-references: GB:AE007869; PIDN:AAK86391.1; PID:G1515521; GSPDB:GN00169
 C:Genetics:
 A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AAPP 4
 ||||
 Db 65 AAPP 68

RESULT 12
 flagellar biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AG2647
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Orage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, M.; Perry, M.; Gordon-Kamm, S.E.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AG2647
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-88 <KUR>
 A:CROSS-references: GB:AE008608; PIDN:ALM1597.1; PID:G17738933; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Flq
 A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AAPP 4
 ||||
 Db 65 AAPP 68

RESULT 13
 S02854
 male accessory gland secretory protein 26Ab precursor - fruit fly (Drosophila melanogaster)
 N:Alternate names: male accessory gland secretory protein mep35b; male paragonial protein
 C:Species: Drosophila melanogaster
 C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Jul-2000
 C:Accession: S02854; S30416; S30420; S30426; S30424
 R:Monma, S.A.; Wolfner, M.F.
 Genes Dev. 2, 1063-1073, 1988
 A:Title: Structure and expression of a Drosophila male accessory gland gene whose product is a secretory protein
 A:Reference number: S02853; MUID:89053045; PMID:3142802
 A:Accession: S02854
 A:Molecule type: DNA
 A:Residues: 1-90 <MON>
 A:CROSS-references: EMBL:Y00219; NID:G8264; PIDN:CAA68367.1; PID:G8266
 A:Experimental source: strain Canton-S
 R:Aguade, M.; Miyashita, N.; Langley, C.H.
 Genetics 132, 755-770, 1992
 A:Title: Polymorphism and divergence in the Mat26A male accessory gland gene region in Drosophila
 A:Reference number: S30407; MUID:93106377; PMID:1361475
 A:Accession: S30408
 A:Molecule type: DNA
 A:Residues: 1-90 <AGU>
 A:CROSS-references: EMBL:X70888; NID:G3402845; PIDN:CAA50233.1; PID:G8234

A:Experimental source: allele NC1
 A:Accession: S30416
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-90 <AGW>
 A:CROSS-references: EMBL:X70892; NID:G8244; PID:G8246
 A:Experimental source: allele NC5
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A:Accession: S30420
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-90 <AGF>
 A:CROSS-references: EMBL:X70894; NID:G8250; PID:G8252
 A:Experimental source: allele NC7
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A:Accession: S30426
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-90 <AGA>
 A:CROSS-references: EMBL:X70897; NID:G8259; PID:G8260
 A:Experimental source: allele NC10
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A:Accession: S30410
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-79, 'V', '81-90 <AG3>
 A:CROSS-references: EMBL:X70899; NID:G3402846; PIDN:CAA50235.1; PID:G8237
 A:Experimental source: allele NC2
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A:Accession: S30424
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-63, 'H', '65-90 <AG4>
 A:CROSS-references: EMBL:X70896; NID:G3406838; PIDN:CAA50249.1; PID:G8258
 A:Experimental source: allele NC9
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 C:Genetics:
 A:Gene: FlyBase:Acp36Ab
 A:CROSS-references: FlyBase:FBgn0002856
 A:Map position: 11/1
 A:Superfamily: male accessory gland secretory protein 26Ab
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:19-90/Product: male accessory gland secretory protein 26Ab #status predicted <MAT>

Query Match 100.0%; Score 21; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AAPP 4
 ||||
 Db 21 AAPP 24

RESULT 14
 S30428
 Met26Ab protein - fruit fly (Drosophila mauritiana)
 C:Species: Drosophila mauritiana
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S30428
 R:Aguade, M.; Miyashita, N.; Langley, C.H.
 Genetics 132, 755-770, 1992
 A:Title: Polymorphism and divergence in the Met26A male accessory gland gene region in Drosophila
 A:Reference number: S30407; MUID:93106377; PMID:1361475
 A:Accession: S30428
 A:Molecule type: DNA
 A:Residues: 1-90 <AGU>
 A:CROSS-references: EMBL:X70898; NID:G8261; PIDN:CAA50253.1; PID:G8263
 C:Genetics:
 A:Gene: FlyBase:Dmau/Acp26Ab
 A:CROSS-references: FlyBase:FBgn0012495
 A:Introns: 11/1

C:Superfamily: male accessory gland secretory protein 26Ab

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
DB 21 AAPF 24

RESULT 15

S30412 Accessory gland secretory protein 26Ab precursor (variant 1) - fruit fly (*Drosophila melanogaster*)
N:Alternate names: male accessory gland secretory protein 355b; Mat26Ab protein
C:Species: *Drosophila melanogaster*
C>Date: 02-Dec-1993 #sequence revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: S30412; S30414; S30418; S30422
R:Aguade, M.; Miyashita, N.; Langley, C.H.
Genetics 132, 765-770, 1992
A:Title: Polymorphism and divergence in the Mat26A male accessory gland gene region in *Drosophila melanogaster*
A:Reference number: S30407; PMID:93106377; PMID:1361475
A:Accession: S30412
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <AGU>
A:Cross-references: EMBL:X70890; NID:G3402847; PIDR:CAM50237.1; PID:G8240
A:Experimental source: allele NC3
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A:Accession: S30414
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <AGU>
A:Cross-references: EMBL:X70891; NID:G8241; PID:G8243
A:Experimental source: allele NC4
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A:Accession: S30418
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <AGF>
A:Cross-references: EMBL:X70893; NID:G8247; PID:G8249
A:Experimental source: allele NC6
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A:Accession: S30422
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-90 <AGU>
A:Cross-references: EMBL:X70895; NID:G8253; PID:G8255
A:Experimental source: allele NC9
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Genetics:
A:Gene: FlyBase:Acp26Aa
A:Cross-references: FlyBase:FBgn0002855
A:Map position: 2
A:Introns: 11/1
C:Superfamily: male accessory gland secretory protein 26Ab
F:1-16/Domain: signal sequence #status predicted <SIG>
F:19-90/Product: male accessory gland secretory protein 26Ab #status predicted <MAT>

Query Match 100.0%; Score 21; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAPF 4
|||
DB 21 AAPF 24

Search completed: February 12, 2003, 10:28:55
Job time : 4.00299 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:

February 12, 2003, 10:04:45 ; Search time 0.537313 Seconds
(without alignments)
308.768 Million cell updates/sec

Title: US-10-036-371-7

Perfect score: 21

Sequence: 1 AAPP 4

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	0	1	SUGA_ACHDO
2	21	100.0	63	1	COXO_BOVIN
3	21	100.0	63	1	COXO_MOUSE
4	21	100.0	90	1	MS2B_DROME
5	21	100.0	90	1	MS2B_DROME
6	21	100.0	90	1	MS2B_DROS
7	21	100.0	90	1	MS2B_DROS
8	21	100.0	102	1	U185_ORYSA
9	21	100.0	102	1	YD86_MYCTU
10	21	100.0	127	1	CAL2_HUMAN
11	21	100.0	127	1	Y355_TREPA
12	21	100.0	128	1	CAL1_HUMAN
13	21	100.0	130	1	CAL2_MOUSE
14	21	100.0	130	1	SZ05_RAT
15	21	100.0	133	1	P232_BOVIN
16	21	100.0	133	1	R8FA_SYNYJ
17	21	100.0	134	1	CH22_RATNO
18	21	100.0	136	1	N12K_AFRSO
19	21	100.0	141	1	CALO_HUMAN
20	21	100.0	141	1	YEF5_YEAST
21	21	100.0	142	1	YEF5_YEAST
22	21	100.0	142	1	PSAH_MAIZE
23	21	100.0	143	1	PSAH_HORVU
24	21	100.0	144	1	PSAH_SPOTL
25	21	100.0	145	1	PSAH_BRAZA
26	21	100.0	145	1	PSH1_ARATH
27	21	100.0	145	1	PSH2_ARATH
28	21	100.0	165	1	VNS3_CVPUF
29	21	100.0	168	1	RBS_SACHY
30	21	100.0	181	1	RBS_JACSA
31	21	100.0	181	1	RBS_JACSA
32	21	100.0	182	1	PAAD_ARCFU
33	21	100.0	186	1	P152_MERTH
34	21	100.0	217	1	YABE_ECOLI
35	21	100.0	200	1	ACD2_EPRHE
36	21	100.0	200	1	R88B_SCHRO
37	21	100.0	200	1	R88B_SCHRO
38	21	100.0	202	1	YD99_HAEIN
39	21	100.0	205	1	YMRP_BACSU
40	21	100.0	206	1	YMR0_YEAST
41	21	100.0	207	1	GTA1_CAEEL
42	21	100.0	208	1	R88_CAEEL
43	21	100.0	209	1	YMR8_BACHD
44	21	100.0	210	1	R88_ARATH
45	21	100.0	217	1	YABE_ECOLI

RESULT 1

SUGA_ACHDO STANDARD; PRT; 5 AA.
AC PI1991;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Suboesophageal ganglion pentapeptide.
OS Acheta domesticus (House cricket).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;
OC Gryllidae; Gryllinae; Acheta.
RX NCBI_TaxID=6997;
RP SEQUENCE.
RA Wicker C.;
RT Isolation and structure of a peptide isolated from the
RT suboesophageal ganglion of Acheta domesticus (orthoptera).
RL Comp. Biochem. Physiol. 88C:185-187(1987).
CC -1- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBESOPHAGEAL
CC GANGLIA.
DR PIR; JS0319; JS0319.
SQ SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 5;
Beet Local Similarity 100.0%; Pred. No. 1.le-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPP 4

Db 2 AAPP 5

RESULT 2

COXO_BOVIN STANDARD; PRT; 63 AA.
AC P00430;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide Viic, mitochondrial precursor
DE (EC 1.9.3.1) (VIIC).
GN COX7C OR COX7CPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC NCBI_TaxID=9913;
RP SEQUENCE FROM N.A.
RX MEDLINE=90045968; PubMed=2554257;
RA Aqua M.S., Lomax M.I., Schon E.A., Grossman L.I.
RT "Nucleotide sequence of a cDNA for bovine cytochrome c oxidase
RT subunit VIIC".
RL Nucleic Acids Res. 17:8376-8376(1989).
RN [2]

RC SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=92009215; PubMed=155579;
 RX Aqua M.S., Bachman N.J., Lomax M.I., Grossman L.I.;
 RT "Characterization and expression of a cDNA specifying subunit VIIC of
 RT bovine cytochrome c oxidase.";
 RL Gene 104:211-217(1991).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9248578; PubMed=9092564;
 RX Seelan R.S., Grossman L.I.;
 RA "Structural organization and promoter analysis of the bovine
 RT cytochrome c oxidase subunit VIIC gene. A functional role for YV1.";
 RL J. Biol. Chem. 272:10175-10181(1997).
 RN [4]
 RN SEQUENCE OF 17-63.
 RC TISSUE=Heart;
 RX MEDLINE=79046803; PubMed=213363;
 RX Buee G., Steffens G.J.;
 RA "Studies on cytochrome c oxidase, II. The chemical constitution of a
 RT short polypeptide from the beef heart enzyme.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 359:1005-1009(1978).
 RN [5]
 RN SEQUENCE OF 17-42.
 RC TISSUE=Liver;
 RX MEDLINE=89000697; PubMed=2844245;
 RX Yanamura M., Zhang Y.-Z., Takamiya S., Capaldi R.A.;
 RA "Tissue-specific differences between heart and liver cytochrome c
 RT oxidase.";
 RL Biochemistry 27:4909-4914(1988).
 RN [6]
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC MEDLINE=96216288; PubMed=8638158;
 RX Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H.,
 RA Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
 RT "The whole structure of the 13-subunit oxidized cytochrome c oxidase
 RT at 2.8 A.";
 RL Science 272:1136-1144(1996).
 RN [7]
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC TISSUE=Heart;
 RX MEDLINE=99190827; PubMed=10089392;
 RX Tomizaki T., Yamashita E., Yamaguchi H., Aoyama H., Tsukihara T.,
 RA Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
 RT "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A
 RT resolution.";
 RL Acta Crystallogr. D 55:31-45(1999).
 RN [8]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RC TISSUE=Heart;
 RX MEDLINE=2035645; PubMed=10711420;
 RX Pei M.J., Yamaguchi H., Inoue N., Yamaguchi H., Tsukihara T.,
 RA Shinzawa-Itoh K., Nakashima R., Yoshikawa S.;
 RT "X-ray structure of fully oxidized cytochrome c oxidase
 RT from bovine heart at 2.9 A resolution.";
 RL Acta Crystallogr. D 56:529-535(2000).
 RN [9]
 RN CHAINS OF CYTOCHROME C OXIDASE. THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE. THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -1- CATALYTIC ACTIVITY: 4 ferricytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -1- TISSUE SPECIFICITY: LIVER, HEART, MUSCLE AND BRAIN. CONTAIN THE
 CC SAME ISOMFORM OF COX VIIC, BUT AT DIFFERENT CONCENTRATIONS.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X15725; CAB57791.1; -;
 DR EMBL; X58823; CAA1627.1; -;
 DR EMBL; U58655; AAC48719.1; -;
 DR PIR; JH0473; OSB08A.
 DR PDB; 1OCC; 07-DEC-96.
 DR PDB; 20CC; 13-JAN-99.
 DR PDB; 1OCR; 29-JUL-99.
 DR PDB; 1OCO; 22-JUL-99.
 DR PDB; 1OCZ; 22-JUL-99.
 DR InterPro; IPR004202; COX7C.
 DR Pfam; PF02935; COX7C; 1.
 KW Oxidoreductase; Mitochondrion; Transit peptide; 3D-structure;
 KW Inner membrane; Transmembrane.
 FT TRANSIT 1 16 MITOCHONDRION.
 FT CHAIN 17 63 CYTOCHROME C OXIDASE POLYPEPTIDE VIIC.
 FT DOMAIN 17 33 MITOCHONDRIAL MATRIX.
 FT TRANSMEM 34 60
 FT DOMAIN 61 63
 FT SEQUENCE 63 AA; 7331 MW; C6A2AD5CC4B63C0B CRC64;
 SO
 Query Match 100.0%; Score 21; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 APPF 4
 Db 50 AAPF 53
 RESULT 3
 COXO_MOUSE STANDARD; PRT; 63 AA.
 ID COXO_MOUSE
 AC P17665; 1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DT 15-JUN-2002 (Rel. 44, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIIC, mitochondrial precursor
 DE (EC 1.3.3.13)
 DE COX7C (Gene)
 DE COX7C (Gene)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 (1)
 PP SEQUENCE FROM N.A.
 RX STRAIN=BA1B/c; TISSUE=Heart;
 RX MEDLINE=90301494; PubMed=2163523;
 RA Akamatsu M., Grossman L.I.;
 RT "Nucleotide sequence of a cDNA for mouse cytochrome c oxidase subunit
 RT VIIC.";
 RL Nucleic Acids Res. 18:3645-3645(1990).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE. THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -1- CATALYTIC ACTIVITY: 4 ferricytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

KW Oxidoreductase; Mitochondrion; Transit peptide.

```

FT TRANSIT 1 16 MITOCHONDRION.
FT CHAIN 17 63
SQ SEQUENCE 63 AA, 7333 MW, C6A2AD4ABEB2C7B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPP 4
   ||||
Db 50 AAPF 53

RESULT 4
MS2B DROMA STANDARD; PRT; 90 AA.
DC P3378; 1994 (Rel. 28, Created)
DT 01-FEB-1994
DT 15-OCT-2001 (Rel. 28, Last annotation update)
DE Accessory gland-specific peptide 26Ab precursor (Male accessory
GLAND secretory protein)
GN ACP26AB OR MST26AB OR MST355B
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7226;
RN SEQUENCE FROM N.A.
RP STRAIN=Robertson C340;
RC MEDLINE=93106377; PubMed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mat26A male accessory gland gene
region in Drosophila.";
RL Mol. Biol. Evol. 15:1040-1046(1998).
RN SEQUENCE FROM N.A.
RP STRAIN=Various strains;
RC MEDLINE=98384843; PubMed=9718731;
RA Taur S.-C., Ting C.-T., Wu C.-I.;
RT Positive selection driving the evolution of a gene of male
production, "Acp26aa, of Drosophila: II. Divergence versus
polymorphism.";
RL Mol. Biol. Evol. 15:1040-1046(1998).
RN SEQUENCE FROM N.A.
RP STRAIN=M036A, M037A, M047A, M079B, M080B, LA25, LA46, and LA108;
RC MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26aa and Acp26Ab
accessory gland genes in the Drosophila melanogaster species
complex.";
RL Genetics 150:1079-1089(1998).
RN SEQUENCE FROM N.A.
RP STRAIN=Robertley; PubMed=10731132;
RA Ananides P., Ginter S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA George R.A., Lewis S.F., Richard S.M., Zhang Q., Chen L.X., Chen B.D.,
RA Suton G.G., Mortman J.F., Yangell M.D., Zhang Q., Chen L.X., Chen B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Belt G.G., Nelson C.R., Miller G.D.,
RA Abril J.F., Asayani A., An H.-J., Andrews-Pfankoch C., Baldwin G.,
RA Balles R.M., Basu A., Bakendell J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Curtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher S., Dahlke C., Davenport L.B., Davies P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Gaster C., Gabriellista C.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Haecker N.L., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hoerlin D., Houlston K., Howard T.C., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karp G., Kravitz S., Li J., Li Z., Lin X.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Li J., Li Z., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusekern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

```

DE Accessory gland-specific peptide 26Ab precursor (Male accessory
GLAND secretory protein)
GN ACP26AB OR MST26AB OR MST355B OR CG9024.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RP STRAIN=Canton-S;
RC MEDLINE=89053045; PubMed=3142802;
RA Nomata S.A., Wolfner M.F.;
RT "Structure and expression of a Drosophila male accessory gland gene
product, Acp26A, encodes a peptide pheromone precursor.";
RL Genes Dev. 2:1163-1073(1988).
RN SEQUENCE FROM N.A.
RP STRAIN=NC1, NC2, NC3, NC4, NC5, NC6, NC7, NC8, NC9, and NC10;
RC MEDLINE=93106377; PubMed=1361475;
RA Aguade M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mat26A male accessory gland gene
region in Drosophila.";
RL Genetics 132:755-770(1992).
RN SEQUENCE FROM N.A.
RP STRAIN=Various strains;
RC MEDLINE=98384843; PubMed=9718731;
RA Taur S.-C., Ting C.-T., Wu C.-I.;
RT Positive selection driving the evolution of a gene of male
production, "Acp26aa, of Drosophila: II. Divergence versus
polymorphism.";
RL Mol. Biol. Evol. 15:1040-1046(1998).
RN SEQUENCE FROM N.A.
RP STRAIN=M036A, M037A, M047A, M079B, M080B, LA25, LA46, and LA108;
RC MEDLINE=99016087; PubMed=9799260;
RA Aguade M.;
RT "Different forces drive the evolution of the Acp26aa and Acp26Ab
accessory gland genes in the Drosophila melanogaster species
complex.";
RL Genetics 150:1079-1089(1998).
RN SEQUENCE FROM N.A.
RP STRAIN=Robertley; PubMed=10731132;
RA Ananides P., Ginter S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA George R.A., Lewis S.F., Richard S.M., Zhang Q., Chen L.X., Chen B.D.,
RA Suton G.G., Mortman J.F., Yangell M.D., Zhang Q., Chen L.X., Chen B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Belt G.G., Nelson C.R., Miller G.D.,
RA Abril J.F., Asayani A., An H.-J., Andrews-Pfankoch C., Baldwin G.,
RA Balles R.M., Basu A., Bakendell J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Curtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher S., Dahlke C., Davenport L.B., Davies P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Gaster C., Gabriellista C.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Haecker N.L., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hoerlin D., Houlston K., Howard T.C., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karp G., Kravitz S., Li J., Li Z., Lin X.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Li J., Li Z., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusekern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

[illegible]

CC entries requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL: X72630; CAA51207.1; -----
CC DR FlyBase: FBgn0012778; Desc\Acp26Ab.
CC KW Behavior: Signal.
CC FT SIGNAL 1 21 POTENTIAL
CC CHAIN 22 90 ACCESSORY GLAND-SPECIFIC PEPTIDE 26AB.
CC SQ SEQUENCE 90 AA: 10270 MW; 99CFCAEE69B7782E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 90;
Best Local Similarity 100.0%; Pred.No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 21 AAPF 24

RESULT 7

MS2B DROS1
ID MS2B DROS1 STANDARD; PRT; 90 AA.

AC P33740;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE Accessory gland-specific peptide 26Ab precursor
DE Accessory gland-specific peptide 26Ab precursor (Male accessory
DE gland secretory protein 355B).
GN ACP26AB OR MST26AB OR MST355B.

OS Drosophila simulans (fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
CC Mecoptera; Ephydroidea; Drosophilidae; Drosophila.
CC NCBI_TaxID=7240;

RP SEQUENCE FROM N.A.
RC STRAIN=PC;
RX MEDL=NC_93106377; PubMed=1361475;

RA Aquada M., Miyashita N., Langley C.H.;
RT "Polymorphism and divergence in the Mt26A male accessory gland gene
RT region in Drosophila.";
RL Genetics 132:755-770(1992).

CC -!- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING
CC MATING AND MAY AFFECT EGG-LAYING AND BEHAVIOR AFTER MATING.
CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLANDS OF MALSS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL: X70899; CAA50255.1; -;
CC DR PIR: S30430; S30430.
CC DR FlyBase: FBgn0012821; Desc\Acp26Ab.
CC KW Behavior: Signal.

FT SIGNAL 1 21 POTENTIAL
CC CHAIN 22 90 ACCESSORY GLAND-SPECIFIC PEPTIDE 26AB.
CC SQ SEQUENCE 90 AA: 10189 MW; F7F8D6E6D04F03E5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 90;
Best Local Similarity 100.0%; Pred.No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 21 AAPF 24

RESULT 8

U185 ORYSA
ID U185 ORYSA STANDARD; PRT; 102 AA.

AC Q94DM8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical protein P0483G10.15.
DE Hypothetical protein P0483G10.15.
OS Oryza sativa (Rice).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC Eriocarpaceae; Oryzae; Oryza.

CC NCBI_TaxID=4530;
RP SEQUENCE FROM N.A.
RC STRAIN=CV; Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0483G10.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE UPF0185 (BM-002) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL: AP003263; BABE3621.1; -;
CC DR InterPro: IP005615; Pfam: PF03671; UPF0185;
CC DR Pfam: PF03671; UPF0185; 1.
CC KW Hypothetical protein.

CC NCBI_TaxID=7240;
RP SEQUENCE 102 AA; 10362 MW; 1E8F85B2F371159 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 102;
Best Local Similarity 100.0%; Pred.No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4
DB 99 AAPF 102

RESULT 9

YD86 MYCTU
ID YD86 MYCTU STANDARD; PRT; 102 AA.

AC P71656;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1386 precursor.
DE RV1386 OR MT1430 OR MTCV21B4.03.
GN Mycobacterium tuberculosis.

CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=1773;
RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;
RX MEDL=U095987; PubMed=9634230;
RA Gordon S.V., Busch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Busch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Badoock K., Bagdasarian K., Gas S., Barry C.E. III, Tekala F.,
RA Badoock K., Bagdasarian K., Gas S., Barry C.E. III, Tekala F.,
RA Horsby T., Jagels K., Knoch A., Fellows T., Chillingworth T., Connor R.,
RA Horsby T., Jagels K., Knoch A., Fellows T., Chillingworth T., Connor R.,
RA Rutter S., Osborne J., Quail M.A., Rajandran S., Rogers J.,
RA Rutter S., Osborne J., Quail M.A., Rajandran S., Rogers J.,
RA Sultana J.E., Taylor K., Skelton S., Smarase S., Gnanapavan K.,
RA Sultana J.E., Taylor K., Skelton S., Smarase S., Gnanapavan K.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).

CC NCBI_TaxID=1773;
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDL=U095987; PubMed=9634230;
RA Gordon S.V., Busch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Busch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Badoock K., Bagdasarian K., Gas S., Barry C.E. III, Tekala F.,
RA Badoock K., Bagdasarian K., Gas S., Barry C.E. III, Tekala F.,
RA Horsby T., Jagels K., Knoch A., Fellows T., Chillingworth T., Connor R.,
RA Horsby T., Jagels K., Knoch A., Fellows T., Chillingworth T., Connor R.,
RA Rutter S., Osborne J., Quail M.A., Rajandran S., Rogers J.,
RA Rutter S., Osborne J., Quail M.A., Rajandran S., Rogers J.,
RA Sultana J.E., Taylor K., Skelton S., Smarase S., Gnanapavan K.,
RA Sultana J.E., Taylor K., Skelton S., Smarase S., Gnanapavan K.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).

QY 1 AAPF 4
DB 99 AAPF 102

```

RP SEQUENCE FROM N.A.
RC STRAIN:GDC 1551 / Oshkosh;
RA Fielemann R., Allard D., Eisen J.A., Carpenter L., White O.,
RA Kotonen J., Dery R., Dodson R., Quinn M.L., Haft D., Hickey E.,
RA Kotonen J.F., Nelson W.C., Umayam L.A., Emolaeva M.D., Salzberg S.L.,
RA Decheray A., Otterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Signal W.,
RA Whole genome comparison of Mycobacterium tuberculosis clinical and
RA laboratory strains",
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2001) to the MYCOBACTERIAL PE FAMILY.
CC - SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z80108; CAB02191.1; -.
CC TIGR; AE007014; AAK45695.1; -.
CC Tuberculin; RV1386; -.
CC InterPro; IPR000084; PE region.
CC Pfam; PF00934; PE; 1.
CC Hypothetical protein; Signal; Complete proteome.
CC SIGNAL 1 31
CC CHAIN 32 102
CC SEQUENCE 102 AA; 9862 MW; 8165F09DB9B9D752 CRC64;
CC
CC Query Match 100.0%; Score 21; DB 1; Length 102;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 AAPP 4
CC 1111
CC 31 AAPF 34
CC
CC RESULT 10
CC CAL2 HUMAN
CC ID CAL2 HUMAN STANDARD; PRT; 127 AA.
CC AC P10092; GUCN9; -.
CC DT 01-MAR-1989 (Rel. 10, Created)
CC DT 01-MAR-1989 (Rel. 10, Last annotation update)
CC DT 21-JUN-2002 (Rel. 41, Last annotation update)
CC DE CGRP; Calcitonin gene-related peptide II precursor (CGRP-II) (Beta-type
CC DE CGRP)
CC OS CAL2 OR CALC2
CC EC 3.4.21.10 (human)
CC EC 3.4.21.10 (human)
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
CC OX NCBI_TaxId=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=87105923; PubMed=3492393;
CC RA Steenbergh P.H., Hoepfner J.W.M., Zandberg J., Visser A.,
CC RA Lips C.J.M., Janz H.S.;
CC RT "Structure and expression of the human calcitonin/CGRP genes.";
CC RL FEBS Lett. 209:197-103(1986).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Brain;
CC RA Strausberg R.;
CC RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC [3]
CC RN SEQUENCE OF 56-127 FROM N.A.
CC RX MEDLINE=85180007; PubMed=2985435;
CC RA Steenbergh P.H., Hoepfner J.W.M., Zandberg J., Lips C.J.M.,
CC RA Janz H.S.;
CC RT "A second human calcitonin/CGRP gene.";
CC RL FEBS Lett. 183:403-407(1985).

```

```

RN [4]
RX PARTIAL SEQUENCE OF 82-108.
RX MEDLINE=87109142; PubMed=3492492;
RA Petermann J.B., Born W., Chang J.Y., Fischer J.A.;
RA "Identification in the human central nervous system, pituitary, and
RA thyroid of a novel calcitonin gene-related peptide, and partial amino
RT acid sequence in the spinal cord.";
RL J. Biol. Chem. 262:542-545(1987).
RN [5]
RP SEQUENCE OF 82-86 AND 104-117.
RP TISSUE=Spinal cord;
RX MEDLINE=90211348; PubMed=2322988;
RX Wimalawansa S.J., Morris H.R., Etienne A., Blench I., Panico M.,
RA McIntyre I.;
RA "Isolation, purification and characterization of beta-hCGRP from
RT human spinal cord.";
RL Biochem. Biophys. Res. Commun. 167:993-1000(1990).
RN [6]
RP SEQUENCE OF 82-104.
RP TISSUE=Phenochromocytoma;
RX MEDLINE=92287083; PubMed=1318039;
RX Kikamura K., Kangawa K., Kawamoto M., Ichiki Y., Matsuo H., Eto T.;
RT "Isolation and characterization of peptides which act on rat
RT platelets from a pheochromocytoma 185-134-141(1992).
CC - FUNCTION: Contributes to vasodilatation. It dilates a variety of
CC vessels including the coronary, cerebral and systemic vasculature.
CC Its abundance in the CNS also points toward a neurotransmitter or
CC inhibitory role.
CC - SUBCELLULAR LOCATION: Secreted.
CC - SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04855; CAC05295.1; -.
CC EMBL; X04857; CAC05295.1; JOINED.
CC EMBL; X04861; CAC05295.1; JOINED.
CC EMBL; BC008428; AAH08428.1; -.
CC EMBL; X02404; CAA26249.1; -.
CC PIR; A25864; A25864.
CC PIR; A34565; A34565.
CC PIR; B26142; B26142.
CC Genew; HGNC:1438; CALCB.
CC MIM; 114160; -.
CC InterPro; IPR001693; Calcitonin-like.
CC InterPro; IPR002163; Calcitonin B.
CC Pfam; PF00214; Calc_CGRP_IAPP; 1.
CC PRINTS; PR00817; CALCITONINB.
CC SMART; SM00113; CALCITONIN; 1.
CC PROSITE; PS00259; CALCITONIN; 1.
CC Cleavage on pair of basic residues; Amidation; Hormone; Signal.
CC SIGNAL 1 25
CC PROPEP 26 79
CC PEPTIDE 82 118
CC PROPEP 124 127
CC DISULFID 83 88
CC MOD RES 118 118
CC CONFLICT 73 73
CC SEQUENCE 127 AA; 13706 MW; 80A71A063CD5ACE7 CRC64;
CC
CC Query Match 100.0%; Score 21; DB 1; Length 127;
CC Best Local Similarity 100.0%; Pred. No. 1.6e+02;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 AAPF 4
CC 1111
CC 25 AAPF 28
CC
CC Db

```

```

RESULT 11
AC 0357 TREPA STANDARD: PRT; 127 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein TP0355.
GN TP0355.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=99332770; PubMed=9665976;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Karsgren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA McDonald L., Natchip, Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Venter J.C., Roberts K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.",
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A2001215; AAC65353.1; -.
DR TIGR; TP0355;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 127 AA; 14526 MW; 1B94BEF0DAC4AD04 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 1,6e-02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
DB 95 AAPF 88
|||||
-----
RESULT 12
CALC_HUMAN STANDARD: PRT; 128 AA.
AC P06881; Q93048; OSUCP06881
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcitonin gene-related peptide I precursor (CGRP-I) (Alpha-type
DE CGRP).
DE CGRP.
GN CALCA OR CALCI.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] -
RP SEQUENCE FROM N.A.
RC MEDLINE=89386053; PubMed=2571128;
RA Bredt P.M., Symes A.J., Thakker R.V., Craig R.K.;
RT "Solution structure of human calcitonin gene-related peptide by 1H
RL NMR and distance geometry with restrained molecular dynamics.",
RL Nucleic Acids Res. 17:6939-7011(1989).
RN [2] -
RP SEQUENCE FROM N.A.
RC MEDLINE=91248117; PubMed=2039456;
RX MEDLINE=85166259; PubMed=3872459;
RA Jonas V., Iin C.B., Kawahara E., Semon D., Swanson L.W.,
RA Mermel J.-J., Evans R.M., Rosenfeld M.G.,
RT "Alternative RNA processing events in human calcitonin/calcitonin
RT gene-related peptide gene expression.",
RN Proc. Natl. Acad. Sci. U.S.A. 82:1994-1998(1985).
RN [3] -
RP SEQUENCE OF 48-119 FROM N.A.
RX MEDLINE=8502523; PubMed=6148938;
RA Nelkin B.D., Rosenfeld K.I., de Bustros A., Leong S.S., Roos B.A.,
RA Baylin S.B.;
RT "Structure and expression of a gene encoding human calcitonin and
RT calcitonin gene related peptide.",
RN Biochem. Biophys. Res. Commun. 123:648-655(1984).
RN [4] -
RP SEQUENCE OF 49-128 FROM N.A.
RX MEDLINE=85230541; PubMed=2408983;
RA Edbrooke M.R., Parker D., Mcvey J.H., Riley J.H., Sorenson G.D.,
RA Pettingill O.S., Craig R.K.;
RT "Expression of the human calcitonin/CGRP gene in lung and thyroid
RT carcinoma.",
RN EMBO J. 4:715-724(1985).
RN [5] -
RP SEQUENCE OF 77-128 FROM N.A.
RC TISSUE=Thyroid carcinoma;
RX MEDLINE=84240176; PubMed=6610687;
RA Steenbergh P.H., Hoppener J.W., Zandberg J., de Ven W.J., Jansz H.S.,
RA Lips C.J.;
RT "Calcitonin gene related peptide coding sequence is conserved in the
RT human genome and is expressed in medullary thyroid carcinoma.",
RN J. Clin. Endocrinol. Metab. 59:358-360(1984).
RN [6] -
RP SEQUENCE OF 77-128 FROM N.A.
RX MEDLINE=87213363; PubMed=3034287;
RA Craig R.K., Riley J.H., Edbrooke M.R., Broad P.M., Poord S.M.,
RA Al-Kawini S.J., Holman J.J., Marshall I.;
RT "Expression and function of the human calcitonin/alpha-CGRP gene in
RT health and disease.",
RN Biochem. Soc. Symp. 52:91-105(1986).
RN [7] -
RP SEQUENCE OF 83-119.
RX MEDLINE=84191466; PubMed=6609312;
RA Morris H.R., Panico M., Etienne T., Tippins J., Girgis S.I.,
RA McIntyre I.;
RT "Isolation and characterization of human calcitonin gene-related
RT peptide and characterization of human calcitonin gene-related
RN peptide 308-746-748(1984).
RN [8] -
RP PARTIAL SEQUENCE OF 83-117
RX MEDLINE=87109142; PubMed=3492492;
RA Petermann J.B., Born W., Chang J.Y., Fischer J.A.;
RT "Identification in the human central nervous system, pituitary, and
RT thyroid of a novel calcitonin gene-related peptide, and partial amino
RT acid sequence in the spinal cord.",
RN J. Biol. Chem. 262:542-545(1987).
RN [9] -
RP SEQUENCE OF 83-108, AND FUNCTION.
RC TISSUE=Phochromocytoma;
RX MEDLINE=92287083; PubMed=1318039;
RA Kitamura K., Kangawa K., Kawamoto M., Ichiki Y., Matsuo H., Eto T.;
RT "Isolation and characterization of peptides which act on rat
RT 1-acetate from a phochromocytoma.",
RN Biochem. Biophys. Res. Commun. 185:134-141(1992).
RN [10] -
RP STRUCTURE BY NMR OF CGRP.
RX MEDLINE=91105142; PubMed=1988044;
RA Breeze A.L., Harvey T.S., Bazzo R., Campbell I.D.;
RT "Solution structure of human calcitonin gene-related peptide by 1H
RT NMR and distance geometry with restrained molecular dynamics.",
RN Biochemistry 30:575-582(1991).
RN [11] -
RP STRUCTURE BY NMR OF CGRP.
RX MEDLINE=91248117; PubMed=2039456;

```

RA Hubbard J.A.M., Martin S.R., Chaplin L.C., Rose C., Kelly S.M.,
 RA "Solution structures of calcitonin-gene-related-peptide analogues of
 RA Calcitonin gene-related peptide and amylin";
 PL Biochem J 275:785-788(1991).
 CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
 CC its vessels including the coronary, cerebral and systemic vasculature.
 CC Its abundance in the CNS also points toward a neurotransmitter or
 CC neuromodulator role. It also elevates platelet CAMP.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (AC P01258), 2 (AC P01258) AND
 CC 3 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X15943; CAA34070.1; -
 DR EMBL; X02330; CAA36190.1; -
 DR EMBL; M12667; AAA51914.1; -
 DR EMBL; M12664; AAA51914.1; JOINED.
 DR EMBL; M12665; AAA51914.1; -
 DR EMBL; M12665; AAA52011.1; -
 DR EMBL; M28637; AAA52012.1; -
 DR EMBL; M26094; AAA51912.1; -
 DR PIR; B22716; TCHUR.
 DR PIR; A22949; A22949.
 DR PIR; A26142; A26142.
 DR PIR; S07644; S07644.
 DR Genew; HGNC:1437; CALCA.
 DR MIM; 114130; -
 DR InterPro; IPR001693; Calcitonin-like.
 DR InterPro; IPR002163; Calcitonin B.
 DR Pfam; PF00214; Calc CGRP_IAPP; 1.
 DR PRINTS; PR00817; CALCITONIN.
 DR SMART; SM00113; CALCITONIN; 1.
 DR PROSITE; PS00258; CALCITONIN; 1.
 DR Cleavage on pair of basic residues; Amidation; Alternative splicing;
 KW Hormone; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 60
 FT PEPTIDE 23 118
 FT PROPEP 23 118
 FT PROPEP 123 169
 FT DISULFID 19 119
 FT DISULFID 19 119
 FT CONFLICT 76 76
 FT CONFLICT 76 76
 FT CONFLICT 76 76
 SQ SEQUENCE 128 AA; 13899 MW; A003A1069260D988 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAPF 4
 DB 25 AAPF 28
 RESULT 13
 CAL2_MOUSE
 ID _CAL2_MOUSE STANDARD; PRT; 130 AA.
 AC Q99MF3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calcitonin gene-related peptide II precursor (CGRP-II) (Beta-type
 DE CGRP).
 GN CALCB.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=21604266; PubMed=11761712;
 RA Thomas P.M., Nasonkin I., Zhang H., Gagel R.F., Cote G.J.;
 RT "Structure of the mouse calcitonin/calcitonin gene-related peptide
 RT alpha and beta genes";
 RL DNA Seq. 12:131-135(2001).
 CC -1- FUNCTION: CGRP induces vasodilatation. It dilates a variety of
 CC vessels including the coronary, cerebral and systemic vasculature.
 CC Its abundance in the CNS also points toward a neurotransmitter or
 CC neuromodulator role (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF355226; AAK16431.1; -
 DR EMBL; AF355224; AAK16431.1; JOINED.
 DR MG0; MG1:415254; CalCDN.
 DR InterPro; IPR000493; Amylin.
 DR InterPro; IPR002163; Calcitonin-like.
 DR PIR; PF00214; Calc CGRP_IAPP; 1.
 DR PRINTS; PR00817; CALCITONIN.
 DR SMART; SM00113; CALCITONIN; 1.
 DR PROSITE; PS00258; CALCITONIN; 1.
 DR Cleavage on pair of basic residues; Amidation; Hormone; Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT PROPEP 27 82 BY SIMILARITY.
 FT PEPTIDE 84 120 CALCITONIN GENE-RELATED PEPTIDE II.
 FT PROPEP 127 130 BY SIMILARITY.
 FT DISULFID 85 90 BY SIMILARITY.
 FT MOD_RES 120 120 AMIDATION (G-121 PROVIDE AMIDE GROUP) (BY
 FT SIMILARITY).
 SQ SEQUENCE 130 AA; 14623 MW; 97299244E8F6C536 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 130;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAPF 4
 DB 26 AAPF 29
 RESULT 14
 SZ05_RAT
 ID _SZ05_RAT STANDARD; PRT; 130 AA.
 AC E97885;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small inducible cytokine B5 precursor (CXCL5) (Cytokine LIX).
 DE SCYB5.
 GN Rattus norvegicus (Rat).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RC Keiner G.S., Maciejewski-Lenoir D., Lee E.D., Maki R.A.;


```

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY PARTICIPATE IN THE RECRUITMENT OF INFLAMMATORY CELLS
CC BY INJURED OR INFECTED TISSUE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; U90448; AAB61460.1; -.
CC HSSP: P10889; IM12.
CC InterPro: IPR001089; CXC chemokine sm1.
CC InterPro: IPR001811; Chemokine IL8.
CC Pfam: PF00048; IL8; 1.
CC PRINTS: PR00437; SMALL_CYTOKINES_CXC.
CC SMART: SM00199; SCV; 1.
CC PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
CC Cytokine; Signal.
CC SIGNAL 1 37 POTENTIAL.
CC FT 1 37 SMALL INDUCIBLE CYTOKINE B5.
CC FT 28 130 BY SIMILARITY.
CC FT DISULFID 16 BY SIMILARITY.
CC FT DISULFID 52 9 BY SIMILARITY.
CC SEQUENCE 130 AA; 14263 MW; C00F6B56052AF4E CRC64;
CC -----
Query Match 100.0%; Score 21; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC 1 AAFP 4
CC 37 AAFP 40
CC Db
CC -----
RESULT 15
ID P212 BOVIN STANDARD; PRT; 133 AA.
DC U9831; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE P2Y purinoceptor 2 (P2Y2) (Fragment).
GN P2RY2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang A.S., Chang S.M., Schilling W.P.;
RT "Identification of multiple P2 purinergic receptor subtypes in bovine
RL aortic endothelial cells (BAECs).";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: receptor for ATP and UTP coupled to G-proteins that
CC (by similar to phosphatidylinositol-calcium second messenger system
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.
CC -----
EMBL; AF005153; AAB62543.1; -.

```

```

DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7cm1;
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2;
DR KEGG: G-protein coupled receptor; Transmembrane.
DR NON_TER 1 1
FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 27 47 4 (POTENTIAL).
FT DOMAIN 48 74 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 75 95 5 (POTENTIAL).
FT DOMAIN 96 116 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 117 133 6 (POTENTIAL).
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 15006 MW; 301B087C62C75B69 CRC64;
CC -----
Query Match 100.0%; Score 21; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC 1 AAFP 4
CC 82 AAFP 85
CC Db
CC -----
Search completed: February 12, 2003, 10:23:20
Job time : 2.53731 secs

```

THIS PAGE BLANK (USPTO)


```

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Alpha-amylase precursor (Fragment).
GN AMYS.
OS Bacillus stearothermophilus.
OC Bacteria, Firmicutes, Bacillus/Clostridium group, Bacillales;
OC Geobacillus.
OK NCBI_TaxID=1422;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91092499; PubMed=2265757;
RA Jorgensen P.L., Hansen C.K., Poulsen G.B., Diderichsen B.;
RT "In vivo genetic engineering: homologous recombination as a tool for
RT plasmid construction.";
RL Gene 96:37-41(1990).
DR EMBL; M62638; AAA2242.1; -.
KW SIGNAL.
FT CHAIN 1 34 POTENTIAL.
FT NON_TER 35 >45 POTENTIAL.
FT 45
FT 45
SQ SEQUENCE 45 AA; 5119 MW; 192F623EA13E6DA6 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 35 AAPF 38

RESULT 3
ID Q9TRE9 PRELIMINARY; PRT; 47 AA.
AC Q9TRE9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome C oxidase subunit VIIC (fragment).
DE Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OK NCBI_TaxID=9823;
RN SEQUENCE
RP MEDLINE=93384597; PubMed=836926;
RA Willard R., Gornwall H., Mutt V.;
RT "Characterization of porcine intestinal cytochrome c oxidase subunit
RT VIIC purified by affinity chromatography."
RL Biochim Biophys Res Commun. 195:746-750(1993).
DR InterPro; IPR004202; COX7C.
DR Pfam; PF02615; COX7C; 1.
SQ SEQUENCE 47 AA; 5457 MW; E11E89CFBC18B5D7 CRC64;

Query Match 100.0%; Score 21; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 34 AAPF 37

RESULT 4
ID Q9TOX9 PRELIMINARY; PRT; 47 AA.
AC Q9TOX9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Deoxyribonuclease F (fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

```

OC Bovidae; Bovinae; Bos.
OK NCBI_TaxID=9913;
[1]
RN SEQUENCE.
RP MEDLINE=94183469; PubMed=8136077;
RA Chang Y.M., Lin S., Liao T.H.;
RT "Bovine pancreatic deoxyribonuclease F: isoelectric focusing, peptide
RT mapping and primary structure.";
RL Biotechnol. Appl. Biochem. 19:129-140(1994).
DR HSP; P00639; 3DN1.
DR InterPro; IPR001582; DNase_I.
DR PRINTS; PR00130; DNASEI.
DR ProDom; PD005408; DNase_I.1.
SQ SEQUENCE 47 AA; 4937 MW; D757A9FA41095799 CRC64;

Query Match 100.0%; Score 21; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 17 AAPF 20

RESULT 5
ID Q8YCK3 PRELIMINARY; PRT; 50 AA.
AC Q8YCK3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein EME110525.
GN EME110525.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OK NCBI_TaxID=29495;
RN SEQUENCE FROM N.A.
RP MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapural V., Redkar R.J., Patra G., Mujir C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Tsalonaki I., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AK009688; AAL51767.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 50 AA; 5208 MW; 8D59A82243EFC17 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 44 AAPF 47

RESULT 6
ID Q9YSN3 PRELIMINARY; PRT; 51 AA.
AC Q9YSN3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Lecithin-cholesterol acyltransferase (EC 2.3.1.43) (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Ou J.-F., Saku K., Liao Y.L., Jini S., Jin W.J., Arakawa K.;
RT "A novel missense variant Thr19Met in the lectin: cholesterol
RT acyltransferase (LCAT) gene associated with fish eye disease."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140503; AAD28484.1; -.
RW Acyltransferase; Transferrase.
RT NCBI_TaxID=9606;
RN [1]
RN SEQUENCE 51 AA; 5666 MW; E8511439A5129965 CRC64;
Query Match 100.0%; Score 21; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
DB 22 AAPF 25

RESULT 7
Q9M7N1 PRELIMINARY; PRT; 57 AA.
AC Q9M7N1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 19, Last annotation update)
DE WAF1 attachment factor 1 (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99418887; PubMed=10489241;
RA Gindullis F., Peffer N.J., Meier I.;
RT "MAF1, a novel plant protein interacting with matrix attachment region
RT binding protein MPF1, is located at the nuclear envelope."
RL Plant Cell 11:1755-1769(1999).
DR EMBL; AF118118; AAF63662.1; -.
RW NON-TER
FT 57
SQ SEQUENCE 57 AA; 6085 MW; E35F7AC99E250F78 CRC64;
Query Match 100.0%; Score 21; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
DB 17 AAPF 20

RESULT 8
Q9ADG2 PRELIMINARY; PRT; 59 AA.
AC Q9ADG2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein SC06101.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-A3(2) / M145;

```

```

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 415:413-417(2002).
RW NCBI_TaxID=14712002;
DR EMBL; AF189148; CAC33946.1; -.
KW Hypothetical protein.
SQ SEQUENCE 59 AA; 6387 MW; 885755656DB6C9EF CRC64;
Query Match 100.0%; Score 21; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
DB 5 AAPF 8

RESULT 9
Q8S4B7 PRELIMINARY; PRT; 69 AA.
AC Q8S4B7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MA-1 kinase (Fragment).
OS Amblystegium humile.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Bryidae; Hypnales; Amblystegiaceae; Amblystegium.
OX NCBI_TaxID=111433;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-AH2C;
RA Vanderpoorten A., Shaw A.J., Cox C.J.;
RT "Reconciled gene trees, molecular evolution, and species relationships
RT in Amblystegium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465020; AAU96441.1; -.
RW Kinase.
FT 69
SQ SEQUENCE 69 AA; 7793 MW; 81C122B4FAC30AC CRC64;
Query Match 100.0%; Score 21; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAPF 4
DB 40 AAPF 43

RESULT 10
Q9DWS4 PRELIMINARY; PRT; 70 AA.
AC Q9DWS4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E1 (GP32) (GP35)]
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX STRAIN=VE3216;
RX MEDLINE=20394384; PubMed=10935993;
RA Argentiní C., Dettori S., Villano V., Guadagnino V., Infantolino D.,
RA Dentico P., Coppola R.C., Rapicetta M.;
RT "Molecular characterisation of HCV genotype 4 isolates circulating in
RT Italy.";
RL J. Med. Virol. 62:84-90(2000).
RC -J. Med. Virol. 62:84-90(2000).
CC -J. Med. Virol. 62:84-90(2000).
CC -J. Med. Virol. 62:84-90(2000).
DR EMBL; A250216; CAC16883.1; -.
DR InterPro; IPR002519; HCV env.
DR Pfam; PF01533; HCV env; 1.
DR Coat protein; Envelope protein; Glycoprotein; Polyprotein;
RW Transmembrane.
FT NON TER 1
FT NON TER 70
FT NON TER 70
SQ SEQUENCE 70 AA; 7698 MW; BD360B781C3BF80F CRC64;

Query Match 100.0%; Score 21; DB 12; Length 70;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 54 AAPF 57

RESULT 11
Q5MWJ2 PRELIMINARY; PRT; 72 AA.
AC Q5MWJ2
DC Q5MWJ2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 21, Last annotation update)
DE MHC class II antigen (fragment).
OS Certhidea olivacea.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidae;
OC Pringillidae; Emberizinae; Certhidea.
OX NCBI_TaxID=46880;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=DI104RF;
RC Sato A., Figueroa F., Mayer W.E., Grant P.R., Grant R., Klein J.;
RT "MHC class II genes of Darwin's finches: divergence by point mutations
RT and reciprocal recombination.";
RL (in) Kasahara M. (eds.);
RL Major histocompatibility complex;
RL evolution, structure, and function, pp.518-541, Springer-Verlag,
RL Tokyo (2000).
DR EMBL; AF164163; AAP36225.1; -.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
FT NON TER 1
FT NON TER 72
FT NON TER 72
SQ SEQUENCE 72 AA; 8760 MW; F62FPF27A7BA672B CRC64;

Query Match 100.0%; Score 21; DB 7; Length 72;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 63 AAPF 66

RESULT 12
Q26512 PRELIMINARY; PRT; 73 AA.
AC Q26512
DC Q26512
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Triosephosphate isomerase (fragment).
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=CHINESE AND PHILIPPINE STRAINS;
RX MEDLINE=97101875; PubMed=8946393;
RA Hooker C.W., Brindley P.J.;
RT "Cloning and characterisation of strain-specific transcripts encoding
RT Triosephosphate isomerase, a candidate vaccine antigen from
RT Schistosoma japonicum.";
RL Mol. Biochem. Parasitol. 82:265-269(1996).
DR EMBL; U57557; AAC47394.1; -.
DR HSBP; P00940; ITIM.
DR InterPro; IPR000652; Triophos_ismrse.
DR Pfam; PF00121; TIM; 1.
DR ProDom; PD001005; Triophos_ismrse; 1.
KW Isomerase.
FT NON TER 1
FT NON TER 73
FT NON TER 73
SQ SEQUENCE 73 AA; 7982 MW; 6B10DB35855F5A33 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 35 AAPF 38

RESULT 13
Q8UEJ5 PRELIMINARY; PRT; 75 AA.
AC Q8UEJ5
DC Q8UEJ5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Atul639.
GN ATU1639
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria, alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN 1
RP SEQUENCE FROM N.A.
RC MEDLINE=11743193;
RC Wood D.W., Satubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr., Grant C.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Palmeri A.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009121; AAL42640.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8397 MW; B3B9D20CE8E82DD CRC64;

Query Match 100.0%; Score 21; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4

```

```

Db      50 AAPF 53
RESULT 14
Q9DFN9
ID      Q9DFN9      PRELIMINARY;      PRT;      76 AA.
AC      Q9DFN9;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Compyltetrahydrofolate dehydrogenase (fragment).
OS      Gillichthys mirabilis (long-jawed mudsucker).
OC      Actinopterygii; Neopterygii; Clupeata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Clupeata; Euteleostomi;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidae;
OC      Gobiidae; Gillichthys.
OK      NCBI_TaxID=8222;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=LIVER;
RX      MEDLINE=21117151; PubMed=11172064;
RA      Gracey A.Y., Troll J.V., Somero G.N.;
RT      "Hypoxia-induced gene expression profiling in the euryoxic fish
RL      Gillichthys mirabilis.";
RL      Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
DR      EMBL; AF266196; AAG1316.1; -.
DR      HSSP; P05091; 1CW3.
DR      InterPro; IPR002086; Aldehyde_dehydr.
DR      Pfam; PF00171; aldedh; 1.
PT      Nucleotide.
SQ      SEQUENCE 76 AA; 8182 MW; 787505971A1ACFA3 CRC64;
Query Match      100.0%; Score 21; DB 13; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPF 4
Db      43 AAPF 46
RESULT 15
Q9BL67
ID      Q9BL67      PRELIMINARY;      PRT;      77 AA.
AC      Q9BL67;
DT      01-OCT-2001 (TrEMBLrel. 18, Created)
DT      01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Hypothetical protein msr1158.
OS      Mesorhizobium loti (Mesorhizobium loti).
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Phyllobacteriaceae; Mesorhizobium.
OK      NCBI_TaxID=381;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MAFF303099;
RX      MEDLINE=21082930; PubMed=11214968;
RA      Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Matsumoto K., Itoh T., Taniguchi M., Nakamura K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
DE      Genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti.
RL      DNA Res. 7:331-338(2000)
DR      EMBL; AP002996; BAB48596.1; -.
PT      Nucleotide.
SQ      SEQUENCE 77 AA; 8189 MW; 8A2BF074E4C16806 CRC64;
Query Match      100.0%; Score 21; DB 16; Length 77;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 AAPF 4
Db      54 AAPF 57
Search completed: February 12, 2003, 10:27:29
Job time : 5.82985 secs
OY      1 AAPF 4
Db      54 AAPF 57

```

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	21	100.0	4	1	US-07-973-335-1	Sequence 1, Appli
2	21	100.0	4	1	US-08-155-331-19	Sequence 19, Appli
3	21	100.0	4	1	US-08-004-643C-2	Sequence 2, Appli
4	21	100.0	4	1	US-08-462-456-5	Sequence 5, Appli
5	21	100.0	4	1	US-08-282-860-1	Sequence 6, Appli
6	21	100.0	4	1	US-07-407-000-6	Sequence 1, Appli
7	21	100.0	4	1	US-08-296-922B-28	Sequence 28, Appli
8	21	100.0	4	1	US-08-345-803A-2	Sequence 2, Appli
9	21	100.0	4	1	US-08-544-143C-9	Sequence 3, Appli
10	21	100.0	4	1	US-08-397-602A-9	Sequence 2, Appli
11	21	100.0	4	1	US-08-437-029-5	Sequence 9, Appli
12	21	100.0	4	1	US-08-424-022-19	Sequence 19, Appli
13	21	100.0	4	1	US-08-439-534-10	Sequence 10, Appli
14	21	100.0	4	1	US-08-434-959-3	Sequence 3, Appli
15	21	100.0	4	1	US-08-434-959-4	Sequence 4, Appli
16	21	100.0	4	1	US-08-434-959-5	Sequence 5, Appli
17	21	100.0	4	1	US-08-434-959-6	Sequence 6, Appli
18	21	100.0	4	1	US-08-434-959-6	Sequence 6, Appli
19	21	100.0	4	1	US-08-460-343B-56	Sequence 56, Appli
20	21	100.0	4	1	US-08-067-180B-2	Sequence 2, Appli
21	21	100.0	4	1	US-08-399-028B-56	Sequence 56, Appli
22	21	100.0	4	1	US-08-470-837-3	Sequence 3, Appli
23	21	100.0	4	2	US-08-722-268-2	Sequence 2, Appli
24	21	100.0	4	2	US-08-504-265B-56	Sequence 56, Appli
25	21	100.0	4	2	US-08-439-005-10	Sequence 10, Appli
26	21	100.0	4	2	US-08-698-575E-2	Sequence 2, Appli
27	21	100.0	4	2	US-08-424-017B-19	Sequence 19, Appli

Qy 1 AAPP 4
 ||||
 Db 1 AAPP 4

RESULT 2

US-08-155-331-19
 ; Sequence 19, Application US/08155331
 ; Patent No. 5449311
 ; GENERAL INFORMATION:
 ; APPLICANT: Forster, Donald C
 ; APPLICANT: Forster, Cindy
 ; APPLICANT: No. 544193119, Kfield
 ; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
 ; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ZymoGenetics, Inc.
 ; STREET: 4225 Roosevelt Way, N.E.
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/155,331
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/985,692
 ; FILING DATE: 02-DEC-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker, Gary E
 ; REGISTRATION NUMBER: 31-684
 ; REFERENCE/DOCKET NUMBER: 92-21C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-547-8080 ext 322
 ; TELEFAX: 206-548-2329
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; FRAGMENT TYPE: peptide
 ; FEATURE KEY: Modified-site
 ; LOCATION: 1..4
 ; OTHER INFORMATION: /label= ALA-1
 ; OTHER INFORMATION: /note= "Amino terminal alanine residue is capped with a succinyl group"

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AAPP 4
 ||||
 Db 1 AAPP 4

RESULT 3

US-08-004-643C-2
 ; Sequence 2, Application US/08004643C
 ; Patent No. 5480779
 ; GENERAL INFORMATION:
 ; APPLICANT: Gunter Fischer & Gerhard K llerz
 ; TITLE OF INVENTION: Cyclosporine Assay
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Gunter Fischer et al. c/o G. P. Katona
 ; STREET: 230 Park Avenue, Room 2200
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10169
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM pc compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: conf. to Patent In Release #1.0, Ver.#1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/004,643C
 ; FILING DATE: 12 January 1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 703,590
 ; FILING DATE: 20 May 1991
 ; APPLICATION NUMBER: 398,082
 ; FILING DATE: 24 August 1989
 ; APPLICATION NUMBER: DD WF 601 F/319 577W
 ; FILING DATE: 07 SEPTEMBER 1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Katona, Gerhard P
 ; REGISTRATION NUMBER: 20, 829
 ; REFERENCE/DOCKET NUMBER: 691-003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 986-3377
 ; TELEFAX: (212) 986-6126
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acid residues
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-004-643C-2

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPP 4
 ||||
 Db 1 AAPP 4

RESULT 4

US-08-462-456-5
 ; Sequence 5, Application US/08462456
 ; Patent No. 5510333
 ; GENERAL INFORMATION:
 ; APPLICANT: Angelastro, Michael R
 ; APPLICANT: Bey, Philippe
 ; APPLICANT: Doherty, Niall S
 ; APPLICANT: Janusz, Michael J
 ; APPLICANT: Mehdi, Shujaath
 ; APPLICANT: Peet, No. 5510333con B
 ; TITLE OF INVENTION: Inhibitors of Cathepsin G and Elastase
 ; TITLE OF INVENTION: for Preventing Connective Tissue Degradation
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marion Merrell Dow Inc.
 ; STREET: 2110 East Galbraith Rd
 ; CITY: Cincinnati P. O. Box 156300

STATE: Ohio
COUNTRY: USA
ZIP: 45215-6300
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,456
FILING DATE: ON: 530
COUNTRY: USA
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/342,999
FILING DATE: US/08/222,552
APPLICATION NUMBER: US/07/987,587
FILING DATE: US/07/704,499
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Nesbitt, Stephen L.
REGISTRATION NUMBER: 28,981
REFERENCE/DOCKET NUMBER: M01593
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 948-7965
TELEFAX: (513) 948-7961
TELEX: 214320
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;
QY 1 AAPP 4
Db 1 AAPP 4
RESULT 5
US-08-282-860-1
Sequence 1, Application US/08282860
Patent No. 5561108
GENERAL INFORMATION:
APPLICANT: Tsay, Grace C.
APPLICANT: Cheung, Neal K. H.
APPLICANT: Brattencourt, Jeffrey D.
TITLE OF INVENTION: Preparation of Alpha-1 Antichymotrypsin
TITLE OF INVENTION: PEPTIDS PROTEINACEOUS DRUG NASAL
TITLE OF INVENTION: POWDER COMPOSITION
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bayer Corporation
STREET: 800 Dwight Way
STREET: P. O. Box 1986
CITY: Berkeley
STATE: California
COUNTRY: USA
ZIP: 94701-1986
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb Storage
COMPUTER: IBM
OPERATING SYSTEM: DOS
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,860
FILING DATE: 07/29/94
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giblin, James A.
REGISTRATION NUMBER: 25772
REFERENCE/DOCKET NUMBER: MSB-7214
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 705-7210
TELEFAX: (510) 705-7304
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: peptide
FEATURE:
NAME/KEY: Modified peptide used for enzyme activity assay
OTHER INFORMATION: peptide is modified with n-terminal
PUBLICATION INFORMATION: eucynyl group and c-terminal pantoimilide group
ADDRESSES: DeWitt, R. G.
AUTHORS: Legendre, C.
AUTHORS: Broderick, J. W.
TITLE: A Sensitive New Substrate for Chymotrypsin
JOURNAL: Analytical Biochemistry
VOLUME: 99
PAGES: 316-329
DATE: 01-NOV-1979
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 4
US-08-282-860-1
Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;
QY 1 AAPP 4
Db 1 AAPP 4
RESULT 6
US-08-407-000-6
Sequence 6, Application US/08407000
Patent No. 5578324
GENERAL INFORMATION:
APPLICANT: Dohi, Masahiko
APPLICANT: Nishibe, Yoshihisa
APPLICANT: Makino, Yuzuki
APPLICANT: Nishibe, Yoshihisa
TITLE OF INVENTION: PEPTIDS PROTEINACEOUS DRUG NASAL
TITLE OF INVENTION: POWDER COMPOSITION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,000
FILING DATE: 29-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/01257
FILING DATE: 29-JUL-1994
CLASSIFICATION: JP-A-5-206922

/ FILING DATE: 30-JUL-1993
 / APPLICATION NUMBER: JP-A-5-235841
 / FILING DATE: 30-AUG-1993
 / APPLICATION NUMBER: JP-A-6-1644
 / FILING DATE: 12-JAN-1994
 / INFORMATION FOR SEQ ID NO: 6:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 4 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: peptide
 / US-08-407-000-6

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 7
 US-07-890-422B-28
 / Sequence 28 Application US/07890422B
 / Patent No. 5602103
 / GENERAL INFORMATION:
 / APPLICANT: THIEL, DWAIN L.
 / APPLICANT: LIPSKY, PETER B.
 / APPLICANT: MCGUIRE, MICHAEL J.
 / TITLE OF INVENTION: DIPEPTIDYL PEPTIDASE-I
 / TITLE OF INVENTION: INHIBITORS AND USES THEREOF
 / NUMBER OF SEQUENCES: 38
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: ARNOLD WHITE & DURKEE
 / STREET: P.O. BOX 4433
 / CITY: HOUSTON
 / STATE: TEXAS
 / COUNTRY: USA
 / ZIP: 77210

/ COMPUTER READABLE FORM:
 / MEDIUM TYPE: FLOPPY DISK
 / COMPUTER: IBM PC COMPATIBLE
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: WORDPERFECT 5.1
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/07/890.422B
 / FILING DATE: 19920529
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: MAYFIELD, DENISE L.
 / REGISTRATION NUMBER: 33,732
 / REFERENCE/DOCKET NUMBER: UTSD:296/MAY
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 512-320-7200
 / TELEFAX: 512-474-7577
 / TELEX: NOT APPLICABLE
 / INFORMATION FOR SEQ ID NO: 28:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 4 amino acid residues
 / TYPE: AMINO ACID
 / TOPOLOGY: linear

US-07-890-422B-28
 Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 8
 US-08-276-936A-2
 / Sequence 2 Application US/08276936A
 / Patent No. 5612194
 / GENERAL INFORMATION:
 / APPLICANT: Harvey Rubin, Barry Cooperman, No. 5612194man Schecter,
 / APPLICANT: Michael Plotkin, Zhi Wang
 / TITLE OF INVENTION: Methods of Producing Effective
 / TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These
 / TITLE OF INVENTION: Inhibitors
 / NUMBER OF SEQUENCES: 30
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Jane Massey Licata, Esq.
 / STREET: 210 Lake Drive East, Suite 201
 / CITY: Cherry Hill
 / STATE: NJ
 / COUNTRY: USA
 / ZIP: 08002

/ COMPUTER READABLE FORM:
 / MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
 / COMPUTER: IBM 486
 / OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 / SOFTWARE: WORDPERFECT 5.1
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/276,936A
 / FILING DATE: July 19, 1994
 / CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 221,078
 / FILING DATE: March 31, 1994
 / APPLICATION NUMBER: 221,171
 / FILING DATE: March 31, 1994
 / APPLICATION NUMBER: 005,908
 / FILING DATE: January 15, 1993
 / APPLICATION NUMBER: 735,335
 / FILING DATE: July 24, 1991

/ ATTORNEY/AGENT INFORMATION:
 / NAME: Jane Massey Licata
 / REGISTRATION NUMBER: 32,257
 / REFERENCE/DOCKET NUMBER: PENN-0027
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (609) 779-8400
 / TELEFAX: (609) 779-8486
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 4
 / TYPE: Amino Acid
 / TOPOLOGY: Linear

US-08-276-936A-2
 Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

RESULT 9
 US-08-345-820B-2
 / Sequence 2 Application US/08345820B
 / Patent No. 5618792
 / GENERAL INFORMATION:
 / APPLICANT:
 / TITLE OF INVENTION: SUBSTITUTED HETEROCYCLIC COMPOUNDS USEFUL AS
 / TITLE OF INVENTION: INHIBITORS OF (SERINE PROTEASES) HUMAN NEUTROPHIL ELASTASE
 / NUMBER OF SEQUENCES: 4
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
 Db 1 AAPF 4

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,820B
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-345-820B-2

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAPF 4
Db 1 AAPF 4

```

```

RESULT 10
US-08-544-143A-3
; Sequence 3, Application US/08544143A
; Patent No. 5646028
; GENERAL INFORMATION:
; APPLICANT: Leigh, Scott D.
; TITLE OF INVENTION: NOVEL ALKALINE PROTEASE AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mayfield, Parsons, Siebert & Heue
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111-4121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,143A
; FILING DATE: 17-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Siebert, Heue & Parsons
; REGISTRATION NUMBER: 37,119
; REFERENCE/DOCKET NUMBER: 0809.054U33
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-362-5556
; TELEFAX: 415-362-5418
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-544-143A-3

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAPF 4
Db 1 AAPF 4

```

```

RESULT 11
US-08-397-602A-9

```

```

; Sequence 9, Application US/08397602A
; Patent No. 5668107
; GENERAL INFORMATION:
; APPLICANT: Wilson, Charles R
; APPLICANT: Tang, Maria R
; APPLICANT: Berger, Harald
; APPLICANT: Christianson, Teresa M
; APPLICANT: Hansen, Dieter
; TITLE OF INVENTION: Expression Systems for the Production
; TITLE OF INVENTION: of Target Proteins in Bacillus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Henkel Corporation Law Department
; STREET: 140 Germantown Pike, Suite 150
; CITY: Plymouth Meeting
; STATE: PA
; COUNTRY: USA
; ZIP: 19362
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,602A
; FILING DATE: 02-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jaeschke, Wayne C
; REGISTRATION NUMBER: 21,062
; REFERENCE/DOCKET NUMBER: D8969/M4828
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 832-2200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note="N-SUCCINYL"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note=
; OTHER INFORMATION: "P-NITROANILIDE"
US-08-397-602A-9

```

```

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAPF 4
Db 1 AAPF 4

```

```

RESULT 12
US-08-437-029-5
; Sequence 5, Application US/08437029
; Patent No. 5668107
; GENERAL INFORMATION:
; APPLICANT: Miller, Edward J.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Inhibiting Elastase
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco

```

```

, STATE: California
, COUNTRY: USA
, ZIP: 94105
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.25
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/437,029
, FILING DATE: 08 May 95
, CLASSIFICATION: 514
, ATTORNEY/AGENT INFORMATION:
, NAME: Dow, Karen B.
, REGISTRATION NUMBER: 29,684
, REFERENCE/DOCKET NUMBER: 017066-000620
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415-326-2400
, TELEFAX: 415-326-2422
, INFORMATION FOR SEQ ID NO: 5:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 4 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
US-08-437-029-5

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 26+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 1 AAPF 4

RESULT 13
US-08-424-022-19
, Sequence 19, Application US/08424022
, Patent No. 5677146
, GENERAL INFORMATION:
, APPLICANT: Foster, Donald C
, APPLICANT: Sprecher, Cindy
, APPLICANT: Roth, 5677146, Kield
, TITLE OF INVENTION: NOVEL HUMAN
, TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
, NUMBER OF SEQUENCES: 19
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: ZymoGenetics, Inc.
, STREET: 1225 Roosevelt Way, N.E.
, CITY: Seattle
, STATE: WA
, COUNTRY: USA
, ZIP: 98105
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.25
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/424,022
, FILING DATE:
, CLASSIFICATION: 514
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 07/985,692
, FILING DATE: 02-DEC-1992
, ATTORNEY/AGENT INFORMATION:
, NAME: Parker, Gary E
, REGISTRATION NUMBER: 31-684
, REFERENCE/DOCKET NUMBER: 92-21C2
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 206-547-8080 ext 322

```

```

, TELEFAX: 206-548-2329
, INFORMATION FOR SEQ ID NO: 19:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 4 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: peptide
, FRAGMENT TYPE: internal
, FEATURES:
, NAME/KEY: Modified-site
, LOCATION: 1..4
, OTHER INFORMATION: /label= ALA-1
, OTHER INFORMATION: /note= "Amino terminal alanine residue is capped
, OTHER INFORMATION: with a succinyl group"
, FEATURES:
, NAME/KEY: Modified-site
, LOCATION: 1..4
, OTHER INFORMATION: /label= Phe-4
, OTHER INFORMATION: /note= "Carboxyl-terminal phenylalanine residue is
, OTHER INFORMATION: capped with p-nitroanil..."
US-08-424-022-19

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 26+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
Db 1 AAPF 4

RESULT 14
US-08-439-534-10
, Sequence 10, Application US/08439534
, Patent No. 5677146
, GENERAL INFORMATION:
, APPLICANT: Lazarus, Robert A.
, APPLICANT: Dennis, Mark S.
, APPLICANT: Ulmer, Jana S.
, TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING
, TITLE OF INVENTION: ECOTIN AND HOMOLOGS THEREOF
, NUMBER OF SEQUENCES: 12
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: Genentech, Inc.
, STREET: 460 Point San Bruno Blvd
, CITY: South San Francisco
, STATE: California
, COUNTRY: USA
, ZIP: 94080
, COMPUTER READABLE FORM:
, MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: patin (Genentech)
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/439,534
, FILING DATE: 11-MAY-1995
, CLASSIFICATION: 435
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/319501
, FILING DATE: 04-OCT-1994
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/121004
, FILING DATE: 14-SEP-1993
, ATTORNEY/AGENT INFORMATION:
, NAME: Kubinec, Jeffrey S.
, REGISTRATION NUMBER: 36,575
, REFERENCE/DOCKET NUMBER: P0859C1D2
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415/952-9881
, TELEFAX: 910/371-7168
, INFORMATION FOR SEQ ID NO: 10:

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-439-534-10

Query Match 100.0%; Score 21; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AAPF 4
 ||||
 Db 1 AAPF 4

Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AAPF 4
 ||||
 Db 1 AAPF 4

Search completed: February 12, 2003, 10:30:11
 Job time : 1.8597 secs

RESULT 15

US-08-434-959-3
 Sequence 3, Application US/08434959

Patent No. 5736520

GENERAL INFORMATION:

APPLICANT: Bey, Philippe
 APPLICANT: Angelastro, Michael R
 APPLICANT: Mehdi, Shujaath
 TITLE OF INVENTION: No. 5736520e1 Peptidase Inhibitors
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marion Merrell Dow Inc.
 STREET: 2110 East Galbraith Rd.
 CITY: Cincinnati P. O. Box 156300
 STATE: Ohio
 COUNTRY: USA

ZIP: 45215-3000

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/434,959

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/750,439

FILING DATE: 20-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/454,803

FILING DATE: 21-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/439,201

FILING DATE: 20-NOV-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/416,817

FILING DATE: 04-OCT-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/254,762

FILING DATE: 07-OCT-1988

ATTORNEY/AGENT INFORMATION:

NAME: Nesbitt, Stephen L

REGISTRATION NUMBER: 28,981

REFERENCE/DOCKET NUMBER: M01368F US

TELECOMMUNICATION INFORMATION:

TELEPHONS: (513) 948-7965

TELEFAX: (513) 948-7961

TELEX: 214320

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-434-959-3

Query Match 100.0%; Score 21; DB 1; Length 4;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:22:36, Search time 0.608955 Seconds
(without alignment)
167.821 Million cell updates/sec

Title: US-10-036-371-7

Perfect score: 21

Sequence: 1 AAPF 4

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB pep.*
2: /cgn2_6/prodata/1/pubaa/PCT_NEW_PUB pep.*
3: /cgn2_6/prodata/1/pubaa/US05_NEW_PUB pep.*
4: /cgn2_6/prodata/1/pubaa/US05_PUBCOMB pep.*
5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB pep.*
6: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB pep.*
7: /cgn2_6/prodata/1/pubaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB pep.*
9: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB pep.*
10: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB pep.*
11: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB pep.*
12: /cgn2_6/prodata/1/pubaa/US10_PUBCOMB pep.*
13: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB pep.*
14: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	9	US-10-041-006A-10
2	21	100.0	4	10	US-09-757-908A-15
3	21	100.0	4	10	US-09-117-3809-2
4	21	100.0	4	10	US-09-935-744-1
5	21	100.0	4	12	US-10-030-624-33
6	21	100.0	4	12	US-10-036-871-7
7	21	100.0	4	12	US-10-033-535-10
8	21	100.0	4	12	US-09-934-927-1
9	21	100.0	6	10	US-10-031-526-4
10	21	100.0	10	9	US-10-068-965-4
11	21	100.0	12	10	US-09-985-157-11
12	21	100.0	25	10	US-09-864-761-38169
13	21	100.0	31	9	US-09-974-879-335
14	21	100.0	45	10	US-09-864-761-48747
15	21	100.0	51	10	US-09-925-297-566
16	21	100.0	62	10	US-09-864-761-42500
17	21	100.0	69	9	US-09-738-626-6222
18	21	100.0	75	10	US-09-864-761-38957
19	21	100.0			

20 21 100.0 85 10 US-09-925-300-1841
21 21 100.0 89 10 US-09-925-301-1206
22 21 100.0 100 10 US-09-893-737-4
23 21 100.0 113 9 US-09-764-668-986
24 21 100.0 125 9 US-09-854-133-59
25 21 100.0 125 9 US-09-738-626-5499
26 21 100.0 125 10 US-09-738-973-59
27 21 100.0 132 10 US-09-738-907-99
28 21 100.0 135 10 US-09-908-322-51
29 21 100.0 137 10 US-09-740-288A-2
30 21 100.0 141 9 US-09-736-457-1822
31 21 100.0 141 9 US-09-902-941-1822
32 21 100.0 141 9 US-09-843-646-1822
33 21 100.0 141 9 US-09-916-780-30
34 21 100.0 148 10 US-09-925-300-1092
35 21 100.0 158 10 US-09-925-301-1170
36 21 100.0 166 10 US-09-739-907-87
37 21 100.0 172 10 US-10-174-590-512
38 21 100.0 178 9 US-10-176-758-512
39 21 100.0 178 9 US-10-175-737-512
40 21 100.0 178 9 US-10-173-706-512
41 21 100.0 178 9 US-10-175-738-512
42 21 100.0 178 9 US-10-175-738-512
43 21 100.0 178 9 US-10-175-752-512
44 21 100.0 178 9 US-10-176-482-512
45 21 100.0 178 9 US-10-176-757-512

ALIGNMENTS

RESULT 1
US-10-041-006A-10
; Sequence 10, Application US/10041.006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041.006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 10
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide substrate
US-10-041-006A-10

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAPF 4

Db 1 AAPF 4

RESULT 2
US-09-757-908A-15
; Sequence 15, Application US/09757908A
; Patent No. US20020052468A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; TITLE OF INVENTION: Disulfide Core Polypeptides
; FILE REFERENCE: 98-13D1
; CURRENT APPLICATION NUMBER: US/09/757.908A
; CURRENT FILING DATE: 2001-01-10


```

; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036,371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-036-371-7

```

```

Query Match          100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPF 4
    ||||
DB 1 AAPF 4

```

```

RESULT 7
US-10-040-655-10
; Sequence 10, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; TITLE OF INVENTION: Protease 1
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/040,655
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide substrate
US-10-040-655-10

```

```

Query Match          100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPF 4
    ||||
DB 1 AAPF 4

```

```

RESULT 8
US-10-033-526-1
; Sequence 1, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE REFERENCE: UCAL217
; CURRENT APPLICATION NUMBER: US/10/033,526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-033-526-1

```

```

Query Match          100.0%; Score 21; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPF 4
    ||||
DB 1 AAPF 4

```

```

RESULT 9
US-09-994-927-1
; Sequence 1, Application US/09994927
; Patent No. US20020127605A1
; GENERAL INFORMATION:
; APPLICANT: Guilford Pharmaceuticals Inc.
; APPLICANT: Hamilton, Gregory
; APPLICANT: Belyakov, Sergei
; APPLICANT: Vaal, Mark
; APPLICANT: Wei, Ling
; APPLICANT: Wu, Yong-Qian
; APPLICANT: Steiner, Joseph
; TITLE OF INVENTION: Bisubstituted Carbocyclic Cyclophilin Binding Compounds and Their
; FILE REFERENCE: 03166.0029.NPUS02
; CURRENT APPLICATION NUMBER: US/09/994,927
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,074
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/291,966
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Model substrate for measuring rotamase inhibition activity
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa at position 1 is N-succinyl
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(6)
; OTHER INFORMATION: Xaa at position 6 is p-nitroanilide
US-09-994-927-1

```

```

Query Match          100.0%; Score 21; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPF 4
    ||||
DB 2 AAPF 5

```

```

RESULT 10
US-10-033-526-4
; Sequence 4, Application US/10033526
; Patent No. US20020147999A1
; GENERAL INFORMATION:
; APPLICANT: Robert W. Mahley
; APPLICANT: Yadong Huang
; TITLE OF INVENTION: Methods of Treating Disorders Related to
; FILE REFERENCE: APOE
; FILE REFERENCE: UCAL217

```

```

; CURRENT APPLICATION NUMBER: US/10/033,526
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,737
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-033-526-4

```

```

Query Match      100.0%; Score 21; DB 12; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPP 4
    ||||
Db 3 AAPP 6

```

```

RESULT 11
US-10-068-965-4
; Sequence 4, Application US/10068965
; Patent No. US20020156123A1
; GENERAL INFORMATION:
; APPLICANT: BALIGNAD, JEAN-LUC
; APPLICANT: PERON, OLIVIER
; TITLE OF INVENTION: NOVEL PHARMACEUTICAL COMPOSITIONS FOR MODULATING
; FILE REFERENCE: DCLERC-2 P1
; CURRENT APPLICATION NUMBER: US/10/068,965
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: PCT/EP00/07731
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 99870171
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caveolin
; OTHER INFORMATION: binding motif
US-10-068-965-4

```

```

Query Match      100.0%; Score 21; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPP 4
    ||||
Db 3 AAPP 6

```

```

RESULT 12
US-09-985-157-11
; Sequence 11, Application US/09985157
; Patent No. US20020146797A1
; GENERAL INFORMATION:
; APPLICANT: SHOKAT, Kevan M.
; TITLE OF INVENTION: Engineered Protein Kinases Which Can Utilize Modified Nucleotide
; FILE REFERENCE: 51538-5002-05
; CURRENT APPLICATION NUMBER: US/09/985,157
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 09/367,065
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US98/02522

```

```

; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/046,727
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: US 08/797,552
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Optimized Ab1 subtrate
US-09-985-157-11

```

```

Query Match      100.0%; Score 21; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AAPP 4
    ||||
Db 5 AAPP 8

```

```

RESULT 13
US-09-864-761-38169
; Sequence 38169, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Remy, Sharon G.
; APPLICANT: Remy, David R.
; APPLICANT: Hunkel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,365
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

```

```

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38169
; TYPE: PRT
; LENGTH: 25
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050342.16
; OTHER INFORMATION: EXPRESSED IN 97474, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.95
US-09-864-761-38169

```

```

Query Match      100.0%; Score 21; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. AC: 48; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

```

```

QY 1 AAPP 4
    |||||
DB 2 AAPF 5

```

```

RESULT 14
US-09-974-879-335
; Sequence 335; Application US/09974879
; Publication No. US20030029003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 335
; LENGTH: 31
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-09-974-879-335
Query Match      100.0%; Score 21; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. AC: 48; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;
QY 1 AAPP 4
    |||||
DB 20 AAPF 23

```

RESULT 15

```

US-09-864-761-48747
; Sequence 48747; Application US/09864761
; Publication No. US020049763A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Rank, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: SeqMax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48747
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006389.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: SWISSPROT HIT: P15772, EVALUJE 8.20e+00
US-09-864-761-48747

```

Query Match 100.0%; Score 21; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAPF 4
|||
Db 39 AAPF 42

Search completed: February 12, 2003, 10:31:07
Job time : 1.60896 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:01:55 ; Search time 13.4328 Seconds
(with alignments)
198.395 Million cell updates/sec

Title: US-10-036-371-1

Perfect score: 95

Sequence: 1 IVGGYXXHSHQVSLNS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	90	94.7	23	21 AAY93937
2	90	94.7	23	22 AAE07943
3	90	94.7	23	22 AAE07944
4	90	94.7	23	20 AAY33347
5	90	94.7	23	20 AAY33347
6	90	94.7	23	21 AAB22966
7	90	94.7	23	21 AAB22966
8	90	94.7	23	21 AAB22966
9	90	94.7	23	21 AAB22966
10	90	94.7	23	21 AAB22966

11	90	94.7	37	23 AAO21360
12	90	94.7	23	22 AAB31579
13	79	83.2	23	20 AAY04644
14	78	82.1	25	21 AAB22966
15	78	82.1	25	21 AAB22966
16	78	82.1	25	21 AAB22966
17	78	82.1	25	21 AAB22966
18	78	82.1	25	21 AAB22966
19	77	81.1	25	23 ABO04645
20	71	74.7	246	19 AAW43260
21	68.5	72.1	19	22 AAB31573
22	66	69.5	20	21 AAY78977
23	66	69.5	246	21 AAY78975
24	66	69.5	247	18 AAW08475
25	65	68.4	240	19 AAW77470
26	65	68.4	247	9 AAF82243
27	65	68.4	247	9 AAF82243
28	64	67.4	224	20 AAR22703
29	64	67.4	225	22 AAB39500
30	64	67.4	230	20 AAB39488
31	64	67.4	241	21 AAB21316
32	64	67.4	247	21 AAB21321
33	64	67.4	247	23 AAU87693
34	64	67.4	254	23 AAU74758
35	63	66.3	20	21 AAB03086
36	63	66.3	20	22 AAB31574
37	63	66.3	223	20 AAB81767
38	63	66.3	223	21 AAY69973
39	63	66.3	224	15 AAK53637
40	63	66.3	229	21 AAY77494
41	63	66.3	230	15 AAK53638
42	63	66.3	231	21 AAY91926
43	63	66.3	231	21 AAY91926
44	63	66.3	233	21 AAY81935
45	62	65.3	223	23 ABB83322

ALIGNMENTS

RESULT 1

ID AAY93937 standard; peptide; 23 AA.
XX
XX AAY93937;
XX
XX
DT 03-OCT-2000 (first entry)
DE N-terminal of an atlantic cod hydrolase enzyme.
XX
XX Transplantation rejection; hydrolase; graft versus host disease;
KW Cell surface adhesion molecule; immune reaction; inflammation; shock;
KW tumour metastasis; autoimmune disease;
KW Krill derived multifunctional enzyme.
XX
XX Gadus sp.
XX
XX W0200038708-A1.
XX
XX 06-JUL-2000.
XX
XX 23-DEC-1999; 99WO-US30818.
XX
XX 24-DEC-1998; 98US-0114147.
XX (PHAI-) PHAIRSON MEDICAL INC.
XX
XX Franklin R., St Pierre V;
XX WPI; 2000-452301/39.
XX
XX Preventing or ameliorating transplantation rejection reactions using
PT hydrolase enzymes

Atlantic cod multi
Amino acid sequenc
Engraulis japonicu
Atlantic cod hydro
Salmon enzyme 1 mu
N-terminal of an a
Engraulis japonicu
Salmon enzyme 1 mu
Human amyloid beta
Amino terminal seq
Canine anionic try
Porcine trypsinogen
Trypsinogen-like p
Human spleen tryps
Human pancreatic t
Human trypsin seri
Human trypsin seri
Human trypsinogen
Human trypsinogen
Human P00000 PPT
N-terminus of bovi
Amino terminal seq
Bovine TRYP peptid
TRYP protein. Bos
Bovine trypsin. B
Bovine trypsinogen
Bovine trypsinogen
Recombinant trypsi
Bovine met-phe-try
Trypsinogen analog
Partial trypsin se

XX Disclosure; Page 26; 66pp; English.

XX The specification describes a method for preventing or ameliorating

XX transplantation rejection reactions for transplantation of immune cells

XX or other tissues. The method comprises treating a source of immune cells

XX with a hydrolase or hydrolase mixture and administering the treated

XX immune cells to a recipient animal. The hydrolase especially has a

XX preference for removing, destroying, inactivating or disabling at least

XX one of CD4, CD8, CD28, ICAM-1, CD54, CD152, an integrin, CD154,

XX and CD90 in contrast to removing, destroying, inactivating or

XX disabling TCR. The methods are useful for preventing graft versus host

XX disease by using hydrolase enzymes to remove the cell surface adhesion

XX molecules which are involved in triggering the immune reactions involved

XX in the diseases. The methods are used for treating or preventing

XX cell-cell or cell-virus adhesion syndrome comprising inflammation, shock,

XX tumour metastases, autoimmune disease, transplantation rejection

XX reactions or microbial infections. The present sequence represents the

XX N-terminal of a hydrolase, which may be used in the method of the

XX invention.

XX Sequence 23 AA;

Query Match 94.7%; Score 90; DB 21; Length 23;

Best Local Similarity 85.0%; Pred. No. 1.2e-08;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQHQVSLNS 20

Db 1 IVGGYECTKHSHQHQVSLNS 20

RESULT 2

AAE07943

ID AAE07943 standard; peptide; 23 AA.

XX AC AAE07943;

XX DT 01-NOV-2001 (first entry)

XX DE N-terminal of atlantic cod I serine multifunctional hydrolase.

XX KW Multifunctional hydrolase; rejection reaction; extra-corporeal;

XX therapy; graft versus host disease; transplantation rejection;

XX autoimmune disease; microbial infection; immune disorder; cytostatic;

XX cystic fibrosis; chronic obstructive pulmonary disease; COPD;

XX atherosclerosis; cancer; asthma; septic shock; toxic shock syndrome;

XX conjunctivitis; reperfusion injury; pain; immunosuppressive;

XX antibacterial; vasotropic; atlantic cod.

XX Gadus morhua I.

XX OS US6232088-B1.

XX PN 15-MAY-2001

XX PD 15-MAY-2001

XX PP 24-DEC-1998; 98US-0220731.

XX PR 08-FEB-1995; 95US-0385540.

XX PR 07-JUN-1995; 95US-0486820.

XX PR 08-FEB-1996; 96US-0600273.

XX PA (PHAI-) PHAIRSON MEDICAL INC.

XX PI Franklin RL, St Pierre Y;

XX DR WPI; 2001-450051/48.

XX Ameliorating, preventing or treating immune rejection reactions, such

XX as graft versus host disease, autoimmune disease, asthma, cancer, by

XX extra-corporeally treating donor tissue with hydrolase such as a

XX protease -

XX Disclosure; Column 15-16; 27pp; English.

PS Disclosure; Column 15-16; 27pp; English.

XX The present invention relates to a method for ameliorating

XX transplantation rejection reaction. The method comprises treating

XX extra-corporeally donor tissue or donor source of immune cells with

XX a rejection ameliorating effective amount of a hydrolase enzyme. The

XX method is useful for ameliorating, treating or preventing immune

XX rejection reactions such as graft versus host diseases, organ or tissue

XX transplantation rejection, autoimmune disease and associated conditions,

XX microbial infection, immune disorder, cystic fibrosis, chronic

XX obstructive pulmonary disease (COPD), atherosclerosis, cancer, asthma,

XX septic shock, toxic shock syndrome, conjunctivitis, reperfusion injury

XX and pain in humans. The present sequence is the N-terminal peptide

XX of atlantic cod I serine multifunctional hydrolase, used in the

XX exemplification of the invention.

XX Sequence 23 AA;

Query Match 94.7%; Score 90; DB 22; Length 23;

Best Local Similarity 85.0%; Pred. No. 1.2e-08;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGYCXHXSHQHQVSLNS 20

Db 1 IVGGYECTKHSHQHQVSLNS 20

RESULT 3

AAE07944

ID AAE07944 standard; peptide; 23 AA.

XX AC AAE07944;

XX DT 01-NOV-2001 (first entry)

XX DE N-terminal of atlantic cod II serine multifunctional hydrolase.

XX KW Multifunctional hydrolase; rejection reaction; extra-corporeal;

XX therapy; graft versus host disease; transplantation rejection;

XX autoimmune disease; microbial infection; immune disorder; cytostatic;

XX cystic fibrosis; chronic obstructive pulmonary disease; COPD;

XX atherosclerosis; cancer; asthma; septic shock; toxic shock syndrome;

XX conjunctivitis; reperfusion injury; pain; immunosuppressive;

XX antibacterial; vasotropic; atlantic cod.

XX Gadus morhua II.

XX OS US6232088-B1.

XX PN 15-MAY-2001

XX PD 15-MAY-2001

XX PP 24-DEC-1998; 98US-0220731.

XX PR 08-FEB-1995; 95US-0385540.

XX PR 07-JUN-1995; 95US-0486820.

XX PR 08-FEB-1996; 96US-0600273.

XX PA (PHAI-) PHAIRSON MEDICAL INC.

XX PI Franklin RL, St Pierre Y;

XX DR WPI; 2001-450051/48.

XX Ameliorating, preventing or treating immune rejection reactions, such

XX as graft versus host disease, autoimmune disease, asthma, cancer, by

XX extra-corporeally treating donor tissue with hydrolase such as a

XX protease -

XX Disclosure; Column 15-16; 27pp; English.

XX The present invention relates to a method for ameliorating

XX transplantation rejection reaction. The method comprises treating

XX extra-corporeally donor tissue or donor source of immune cells with

CC a rejection ameliorating effective amount of a hydrolase enzyme. The
 CC method is useful for ameliorating, treating or preventing immune
 CC rejection reactions such as graft versus host diseases, organ or tissue
 CC microcirculation rejection, autoimmune disease and associated conditions,
 CC obstructive pulmonary disease (COPD), attherosclerosis, cancer, asthma,
 CC septic shock, toxic epidermal necrolysis, conjunctivitis, reperfusion injury
 CC and pain in humans. The present invention is the N-terminal peptide
 CC of atlantic cod II serine multifunctional hydrolase, used in the
 CC exemplification of the invention.
 CC Note: The sequence is stated as being the same as that shown as SEQ
 CC ID NO 19 (see AAE07951) in column 31-42 of the specification. However
 CC the sequences differ at several positions.

XX
 XX
 SQ Sequence 23 AA;
 Query Match 94.7%; Score 90; DB 22; Length 23;
 Best Local Similarity 85.0%; Pred. No. 1.2e-08;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 IVGGYXCXHXSHQHQVSLNS 20
 Db 1 IVGGYECTKHSQHQVSLNS 20
 ||||| | ||||| |||||

RESULT 4
 AAY33347
 ID AAY33347 standard; peptide; 25 AA.
 XX
 XX AAY33347;
 AC
 XX
 XX
 XX 29-NOV-1999 (first entry)
 XX
 XX Atlantic cod hydrolase N-terminal fragment II.
 XX
 KW Multifunctional enzyme; treatment; acne; eczema; chymotrypsin; trypsin;
 KW collagenase; elastase; exopeptidase; cell surface receptor; anti-acne;
 KW antiseborrheic; dermatological; anti-eczema; adhesion molecule; ICAM-1;
 KW CD54; ICAM-2; VCAM-1; CD4; CD8; CD28; CD29D; CD31; CD44; CD49; CD62L;
 KW CD102; GM1 ceramide; krill; atlantic cod.
 OS
 OS Gadus SP.
 XX
 XX US5958406-A.

XX
 XX
 XX 28-SEP-1999
 XX
 XX 08-FEB-1996; 96US-0600273.
 XX
 XX 22-NOV-1994; 94US-0388501.
 XX 08-FEB-1995; 95US-0385540.
 XX 07-JUN-1995; 95US-0486820.
 XX (PHAI-) PHAIRSON MEDICAL INC.
 XX
 XX Lindblom R, Kay J, Franklin RL, De Faire JR;
 XX WPI; 1999-561004/47.
 XX
 XX Treating acne and eczema using a krill-derived multifunctional enzyme -
 XX Dislosure; Column 21-22; 42pp; English.

XX
 XX This invention describes a novel method for treating acne and eczema
 XX using a krill-derived multifunctional enzyme (I) which comprises 2 or
 XX more of the activities of chymotrypsin, trypsin, collagenase, elastase
 XX or exopeptidase and is reactive with cell surface receptors such as
 XX proteins or glycoproteins. The product of the invention have
 XX antiseborrheic, anti-acne, dermatological and anti-eczema activity.
 XX (I) removes or inactivates cell surface receptors (proteins and
 XX glycoproteins) and adhesion molecules such as ICAM-1 (i.e. CD54), ICAM-2,
 XX VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD44, CD49, CD62L CD102 and the
 XX asialo GM1 ceramide. This sequence represents the N-terminal fragment of

CC an atlantic cod (Gadus sp.) hydrolase which is used to describe the
 CC method of the invention.
 XX
 XX
 SQ Sequence 25 AA;
 Query Match 94.7%; Score 90; DB 20; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e-08;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 IVGGYXCXHXSHQHQVSLNS 20
 Db 1 IVGGYECTKHSQHQVSLNS 20
 ||||| | ||||| |||||

RESULT 5
 AAY33348
 ID AAY33348 standard; peptide; 25 AA.
 XX
 XX AAY33348;
 AC
 XX 29-NOV-1999 (first entry)
 XX
 XX Krill-derived multifunctional enzyme N-terminal fragment 3.
 XX
 KW Multifunctional enzyme; treatment; acne; eczema; chymotrypsin; trypsin;
 KW collagenase; elastase; exopeptidase; cell surface receptor; anti-acne;
 KW antiseborrheic; dermatological; anti-eczema; adhesion molecule; ICAM-1;
 KW CD54; ICAM-2; VCAM-1; CD4; CD8; CD28; CD29D; CD31; CD44; CD49; CD62L;
 KW CD102; GM1 ceramide; krill.
 OS
 OS Euphasia sp.
 XX
 XX US5958406-A.
 XX
 XX 28-SEP-1999.
 XX
 XX 08-FEB-1996; 96US-0600273.
 XX
 XX 22-NOV-1994; 94US-0388501.
 XX 08-FEB-1995; 95US-0385540.
 XX 07-JUN-1995; 95US-0486820.
 XX (PHAI-) PHAIRSON MEDICAL INC.
 XX
 XX Lindblom R, Kay J, Franklin RL, De Faire JR;
 XX WPI; 1999-561004/47.

XX
 XX Treating acne and eczema using a krill-derived multifunctional enzyme -
 XX Claim 1; Column 57-58; 42pp; English.
 XX
 CC This invention describes a novel method for treating acne and eczema
 CC using a krill-derived multifunctional enzyme (I) which comprises 2 or
 CC more of the activities of chymotrypsin, trypsin, collagenase, elastase
 CC or exopeptidase and is reactive with cell surface receptors such as
 CC proteins or glycoproteins. The product of the invention have
 CC antiseborrheic, anti-acne, dermatological and anti-eczema activity.
 CC (I) removes or inactivates cell surface receptors (proteins and
 CC glycoproteins) and adhesion molecules such as ICAM-1 (i.e. CD54), ICAM-2,
 CC VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD44, CD49, CD62L CD102 and the
 CC asialo GM1 ceramide. This invention describes the N-terminal fragment of
 CC a krill multifunctional enzyme which is used in the method of the
 CC invention.

XX
 XX
 SQ Sequence 25 AA;
 Query Match 94.7%; Score 90; DB 20; Length 25;
 Best Local Similarity 85.0%; Pred. No. 1.3e-08;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 IVGGYXCXHXSHQHQVSLNS 20
 ||||| | ||||| |||||

Db 1 IVGGYECTRHSHQHQVSLNS 20

RESULT 6

AA222966

ID AA222966 standard; peptide, 25 AA.

XX AA222966;

XX 10-JAN-2001 (first entry)

XX Atlantic cod multifunctional hydrolase I peptide.

DE Multifunctional hydrolase; multifunctional protein;

KW protease activity; chymotrypsin; trypsin; collagenase;

KW elastase; wound healing; corneal ulcer; internal trauma;

KW internal surgical wound; nerve; tendon; sheath;

KW adhesion formation inhibition; ophthalmological; vulnery.

XX Gadus morhua.

OS WO200049991-A2.

PV 31-AUG-2000.

PD 29-JUN-1999; 99WO-US14751.

XX 23-FEB-1999; 99US-0256484.

PR (PHAI-) PHAIRSON MEDICAL INC.

XX Franklin R, Cowling D, Hubbell JA;

PI WPI; 2000-587120/55.

DR Use of microgel comprising crosslinked polyanionic polymer and

XX optionally proteolytic enzyme in treatment of corneal wounds e.g.

PT ulcers and abrasions and internal trauma e.g. surgical wounds and for

PT treating implants to reduce adhesions -

XX Disclosure; Page 25; 85pp; English.

XX The invention relates to use of a microgel comprising a crosslinked

CC polyanionic polymer in compositions for treatment of an area affected by

CC a trauma selected from corneal wounds and internal trauma. The

CC compositions further comprise a multifunctional protease having

CC activities selected from two of chymotrypsin activity, trypsin activity,

CC collagenase activity and elastase activity. In particular, the protease

CC may be one of six isoforms of white shrimp (Penaeus vannamei)

CC multifunctional hydrolase (AA222947-B22952), or a multifunctional enzyme

CC comprising one of the peptide sequences given in AA222953-B22967.

CC (including infected ulcers), or abrasions or a chemical or physical

CC insult to the cornea that are likely to give rise to a corneal ulcer.

CC They are also used to treat internal traumas selected from an

CC internal surgical wound, or a trauma to a membrane that covers either

CC an internal organ or tissue or the cavity in which one or more internal

CC organs or tissues reside, with the membrane being selected from the

CC peritoneum, the pericardium, the epicardium and the pleura and a

CC or a nerve sheath, where the trauma is susceptible to giving rise to

CC adhesions. The amount of microgel administered is sufficient to prevent

CC or reduce formation or reformation of adhesions, and the composition can

CC also be used to treat an implant to reduce formation of adhesions.

CC Sequences AA222953-B22967 represent peptide sequences from a variety of

CC multifunctional hydrolases. Multifunctional hydrolases comprising one of

CC these sequences are claimed for use in compositions of the invention.

XX Sequence 25 AA;

SQ Query Match 94.7%; Score 90; DB 21; Length 25;

Best Local Similarity 85.0%; Pred. No. 1.3e-06;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

^ . . .

QY 1 IVGGYCKXKSHQHQVSLNS 20

DB 1 IVGGYECTRHSHQHQVSLNS 20

RESULT 7

AA222967

ID AA222967 standard; peptide, 25 AA.

XX AA222967;

XX 10-JAN-2001 (first entry)

XX Atlantic cod multifunctional hydrolase II peptide.

DE Multifunctional hydrolase; multifunctional protein;

KW protease activity; chymotrypsin; trypsin; collagenase;

KW elastase; wound healing; corneal ulcer; internal trauma;

KW internal surgical wound; nerve; tendon; sheath;

KW adhesion formation inhibition; ophthalmological; vulnery.

XX Gadus morhua.

OS WO200049991-A2.

PV 31-AUG-2000.

PD 29-JUN-1999; 99WO-US14751.

XX 23-FEB-1999; 99US-0256484.

PR (PHAI-) PHAIRSON MEDICAL INC.

XX Franklin R, Cowling D, Hubbell JA;

PI WPI; 2000-587120/55.

DR Use of microgel comprising crosslinked polyanionic polymer and

XX optionally proteolytic enzyme in treatment of corneal wounds e.g.

PT ulcers and abrasions and internal trauma e.g. surgical wounds and for

PT treating implants to reduce adhesions -

XX Disclosure; Page 25; 85pp; English.

XX The invention relates to use of a microgel comprising a crosslinked

CC polyanionic polymer in compositions for treatment of an area affected by

CC a trauma selected from corneal wounds and internal trauma. The

CC compositions further comprise a multifunctional protease having

CC activities selected from two of chymotrypsin activity, trypsin activity,

CC collagenase activity and elastase activity. In particular, the protease

CC may be one of six isoforms of white shrimp (Penaeus vannamei)

CC multifunctional hydrolase (AA222947-B22952), or a multifunctional enzyme

CC comprising one of the peptide sequences given in AA222953-B22967.

CC (including infected ulcers), or abrasions or a chemical or physical

CC insult to the cornea that are likely to give rise to a corneal ulcer.

CC They are also used to treat internal traumas selected from an

CC internal surgical wound, or a trauma to a membrane that covers either

CC an internal organ or tissue or the cavity in which one or more internal

CC organs or tissues reside, with the membrane being selected from the

CC peritoneum, the pericardium, the epicardium and the pleura and a

CC or a nerve sheath, where the trauma is susceptible to giving rise to

CC adhesions. The amount of microgel administered is sufficient to prevent

CC or reduce formation or reformation of adhesions, and the composition can

CC also be used to treat an implant to reduce formation of adhesions.

CC Sequences AA222953-B22967 represent peptide sequences from a variety of

CC multifunctional hydrolases. Multifunctional hydrolases comprising one of

CC these sequences are claimed for use in compositions of the invention.

XX Sequence 25 AA;

SQ Query Match 94.7%; Score 90; DB 21; Length 25;

Best Local Similarity 85.0%; Pred. No. 1.3e-06;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Query Match          94.7%; Score 90; DB 21; Length 25;
Best Local Similarity 85.0%; Pred. No. 1.3e-08;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYCKXHSQAHQVSLNS 20
DB 1 IVGGYECTHSQAHQVSLNS 20

RESULT 8
AAO21358
ID AAO21358 standard; Peptide; 25 AA.
AC AAO21358;
XX
XX
DT 05-AUG-2002 (first entry)
DE Atlantic cod I multifunctional hydrolase peptide region.
XX
XX Tranquiliser; vulnery; cytostatic; gynaecological; antiinflammatory;
XX corneal wound; internal trauma; polyanionic polymer; corneal ulcer;
XX corneal abrasion; surgical wound; organ membrane trauma; tissue; cavity;
XX adhesion; cervical spondylosis; cumulative trauma disorder; appendicitis;
XX endometriosis; pelvic inflammatory disease; adhesive peritonitis;
XX peridentitis; pericarditis; pleuritis; inflammatory disease;
XX multifunctional protein; hydrolase; enzyme; atlantic cod.
XX
XX Gadus morhua.
OS
XX WO200215913-A1.
XX
XX 28-FEB-2002.
PD
XX 23-AUG-2000; 2000WO-US23072.
XX
XX 23-AUG-2000; 2000WO-US23072.
XX
XX (PHAI-) PHAIRSON MEDICAL INC.
XX (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
XX
XX Franklin R, Cowling DSP, Hubbell JA, Van De Wetering P;
XX WPI; 2002-425632/45.
XX
XX Treatment of corneal wounds, internal trauma or inflammatory diseases
XX comprises administration of a polyanionic polymer.
XX
XX Disclosure; Page 24; 59pp; English.
XX
XX The invention relates to a method for treating corneal wounds or internal
XX trauma. The method provides administration of a composition comprising a
XX polyanionic polymer. The method is useful for treating corneal wounds,
XX including corneal ulcers, corneal abrasions or injuries which may give
XX rise to corneal ulcers. The method is also useful for treating internal
XX trauma, including surgical wounds, trauma to a membrane that covers an
XX organ, tissue or cavity or wounds, trauma to a membrane that covers an
XX further method is useful for treating cervical spondylosis, cumulative
XX trauma disorder, endometriosis, pelvic inflammatory disease, adhesive
XX peritonitis, appendicitis, peridentitis, pericarditis and pleuritis or
XX other inflammatory diseases. This sequence represents the atlantic cod I
XX multifunctional hydrolase peptide region relating to the polyanionic
XX polymers of the invention.
XX
XX Sequence 25 AA;
QY Query Match          94.7%; Score 90; DB 23; Length 25;
   Best Local Similarity 85.0%; Pred. No. 1.3e-08;
   Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYCKXHSQAHQVSLNS 20
DB 1 IVGGYECTHSQAHQVSLNS 20

RESULT 10
AAAY93946
ID AAY93946 standard; Peptide; 37 AA.
XX
XX AAY93946;
XX
XX 03-OCT-2000 (first entry)

```

```

RESULT 9
AAO21359
ID AAO21359 standard; Peptide; 25 AA.
XX
XX AAO21359;
XX
XX 05-AUG-2002 (first entry)
XX
XX Atlantic cod II multifunctional hydrolase peptide region.
XX
XX Tranquiliser; vulnery; cytostatic; gynaecological; antiinflammatory;
XX corneal wound; internal trauma; polyanionic polymer; corneal ulcer;
XX corneal abrasion; surgical wound; organ membrane trauma; tissue; cavity;
XX adhesion; cervical spondylosis; cumulative trauma disorder; appendicitis;
XX endometriosis; pelvic inflammatory disease; adhesive peritonitis;
XX peridentitis; pericarditis; pleuritis; inflammatory disease;
XX multifunctional protein; hydrolase; enzyme; atlantic cod.
XX
XX Gadus morhua.
OS
XX WO200215913-A1.
XX
XX 28-FEB-2002.
PD
XX 23-AUG-2000; 2000WO-US23072.
XX
XX 23-AUG-2000; 2000WO-US23072.
XX
XX (PHAI-) PHAIRSON MEDICAL INC.
XX (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
XX
XX Franklin R, Cowling DSP, Hubbell JA, Van De Wetering P;
XX WPI; 2002-425632/45.
XX
XX Treatment of corneal wounds, internal trauma or inflammatory diseases
XX comprises administration of a polyanionic polymer.
XX
XX Disclosure; Page 24; 59pp; English.
XX
XX The invention relates to a method for treating corneal wounds or internal
XX trauma. The method provides administration of a composition comprising a
XX polyanionic polymer. The method is useful for treating corneal wounds,
XX including corneal ulcers, corneal abrasions or injuries which may give
XX rise to corneal ulcers. The method is also useful for treating internal
XX trauma, including surgical wounds, trauma to a membrane that covers an
XX organ, tissue or cavity or wounds, trauma to a membrane that covers an
XX further method is useful for treating cervical spondylosis, cumulative
XX trauma disorder, endometriosis, pelvic inflammatory disease, adhesive
XX peritonitis, appendicitis, peridentitis, pericarditis and pleuritis or
XX other inflammatory diseases. This sequence represents the atlantic cod II
XX multifunctional hydrolase peptide region relating to the polyanionic
XX polymers of the invention.
XX
XX Sequence 25 AA;
QY Query Match          94.7%; Score 90; DB 23; Length 25;
   Best Local Similarity 85.0%; Pred. No. 1.3e-08;
   Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVGGYCKXHSQAHQVSLNS 20
DB 1 IVGGYECTHSQAHQVSLNS 20

RESULT 10
AAAY93946
ID AAY93946 standard; Peptide; 37 AA.
XX
XX AAY93946;
XX
XX 03-OCT-2000 (first entry)

```

XX N-terminal of an atlantic cod trypsin hydrolase enzyme.
 DE Transplantation rejection; hydrolase; graft versus host disease;
 KW cell surface adhesion molecule; immune reaction; inflammation; shock;
 KW tumour metastasis; autoimmune disease,
 KW krill derived multifunctional enzyme.
 OS Gadus sp.
 XX Key Location/Qualifiers
 PH Misc-difference 5 /note= "any amino acid"
 PT WO200039708-A1.
 XX 06-JUL-2000.
 XX 23-DEC-1999; 99WO-US30818.
 XX 24-DEC-1998; 98US-0114147.
 XX (PHAI-) PHAIRSON MEDICAL INC.
 PA Franklin RL, St Pierre Y;
 PT WPI, 2000-452301/39.
 XX Preventing or ameliorating transplantation rejection reactions using
 PT hydrolase enzymes -
 PS Disclosure; Page 26; 66pp; English.
 XX The specification describes a method for preventing or ameliorating
 CC transplantation rejection reactions for transplantation of immune cells
 CC or other tissues. The method comprises treating a source of immune cells
 CC with a hydrolase or hydrolase mixture and administering the treated
 CC immune cells to a recipient animal. The hydrolase especially has a
 CC preference for removing, destroying, inactivating, or disabling at least
 CC one of CD4, CD8, CD28, ICAM-1, CD54, CD59, an integrin, CD154,
 CC CD40 and CD80 in contrast to removing, for preventing, inactivating or
 CC disabling TCR. The methods are useful for preventing graft versus host
 CC disease by using hydrolase enzymes to remove the cell surface adhesion
 CC molecules which are involved in triggering the immune reactions involved
 CC in the diseases. The methods are used for treating or preventing
 CC cell-cell or cell-vacuole adhesion syndrome comprising inflammation, shock,
 CC tumour metastasis, autoimmune disease, transplantation rejection
 CC N-terminal of a trypsin hydrolase, which may be used in the method of
 CC the invention.
 XX Sequence 37 AA;
 SQ Query Match 94.7%; Score 90; DB 21; Length 37;
 Best Local Similarity 85.0%; Pred. No. 2e-08;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IVGGYCXHXSHQHQVSLNS 20
 DB 1 IVGGYQCEAHSHQHQVSLNS 20
 RESULT 11
 ID AAO21360 standard; Peptide; 37 AA.
 XX AAO21360;
 AC AAO21360;
 XX 05-AUG-2002 (first entry)
 DT Atlantic cod multifunctional hydrolase N-terminal peptide region.
 DE Tranquilliser; vulnerary; cyrostatic; gynaecological; antiinflammatory;
 KW

KW corneal wound; internal trauma; polyanionic polymer; corneal ulcer;
 KW abrasion; surgical wound; organ membrane tissue; cavity;
 KW adhesion; cervical spondylosis; cumulative trauma disorder; appendicitis;
 KW endometriosis; pelvic inflammatory disease; adhesive peritonitis;
 KW peridontitis; pericarditis; pleuritis; inflammatory disease;
 KW multifunctional protein; hydrolase; enzyme; atlantic cod.
 OS Gadus morhua.
 XX WO200215913-A1.
 PN 28-FEB-2002.
 XX 23-AUG-2000; 2000WO-US23072.
 PF 23-AUG-2000; 2000WO-US23072.
 PR (PHAI-) PHAIRSON MEDICAL INC.
 PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
 XX Franklin R, Cowling DSP, Hubbell JA, Van De Wetering P;
 XX WPI; 2002-425632/45.
 XX Treatment of corneal wounds, internal trauma or inflammatory diseases
 PT comprises administration of a polyanionic polymer -
 PS Disclosure; Page 24; 59pp; English.
 XX The invention relates to a method for treating corneal wounds or internal
 CC trauma. The method provides administration of a composition comprising a
 CC polyanionic polymer. The method is useful for treating corneal wounds,
 CC including corneal ulcers, corneal abrasions or injuries which may give
 CC rise to corneal ulcers. The method is also useful for treating internal
 CC trauma, including surgical wounds, trauma to a membrane that covers an
 CC organ, tissue or cavity or trauma likely to give rise to adhesions.
 CC further method is useful for treating cervical spondylosis, adhesive
 CC trauma disorder, endometriosis, pelvic inflammatory disease, cumulative
 CC pericarditis, appendicitis, peridontitis, pericarditis and pleuritis or
 CC other inflammatory diseases. This sequence represents an N-terminal
 CC peptide region of an atlantic cod multifunctional hydrolase relating to
 CC the polyanionic polymers of the invention.
 XX Sequence 37 AA;
 SQ Query Match 94.7%; Score 90; DB 23; Length 37;
 Best Local Similarity 85.0%; Pred. No. 2e-08;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IVGGYCXHXSHQHQVSLNS 20
 DB 1 IVGGYQCEAHSHQHQVSLNS 20
 RESULT 12
 ID AAB31579 standard; peptide; 225 AA.
 XX AAB31579;
 AC AAB31579;
 XX 20-APR-2001 (first entry)
 DE Amino acid sequence of cod trypsin isozymes.
 XX Fish; serine proteinase; pain; acute inflammation; chronic inflammation;
 KW arthritis; inflamed joint; bursitis; osteoarthritis; septic arthritis;
 KW rheumatoid arthritis; juvenile rheumatoid arthritis; fibromyalgia;
 KW systemic lupus erythematosus; phlebitis; tendinitis; psoriasis;
 KW acne; eczema; facial seborrheic eczema; foreknin infection;
 KW athlete's foot; fistulae infection; ulcer; navel infection; wrinkle;
 KW scar; keloid; boil; wart; allergic itch; hemorrhoid; wound;
 KW fungal infection; autoimmune disease.
 XX

OS	Gadus callarias.	
XX		
PH	Key	Location/Qualifiers
FT	Misc-difference 9	/label= Lys, Arg
FT	Misc-difference 24	/label= Tyr, Phe
FT	Misc-difference 32	/label= Lys, Gln
FT	Misc-difference 33	/label= Asp, Gln
FT	Misc-difference 64	/label= Tyr, Phe
FT	Misc-difference 71	/label= Ile, Xaa
FT	Misc-difference 82	/note= "Xaa is any amino acid or absent"
FT	Misc-difference 92	/label= Asp, Asn
FT	Misc-difference 129	/label= Lys, Glu
FT	Misc-difference 136	/label= Ala, Asp
FT	Misc-difference 138	/label= Val, Cys
FT	Misc-difference 204	/label= Asn, Ser
FT	Misc-difference 712	/label= His, Asn
FT	Misc-difference 771	/label= Met, Val, Cys
FT	Misc-difference 224	/label= Asn, Ser
XX	W0200078332-A2.	
XX		
PD	28-DEC-2000.	
XX		
XX	15-JUN-2000, 2000WO-1S00005.	
PR	18-JUN-1999; 99IS-0005086.	
PA	(BJAR)/ BJARNASON J B.	
XX		
PI	Bjarnason JB;	
XX		
DR	WPI: 2001-091493/10.	
XX		
PT	Fish serine proteinase, useful as a cosmetic, medicament for treating eczema, psoriasis, arthritis, and in the manufacture of the medicament for treating, preventing pathogenic diseases involving receptor mediated binding	
PT		
PT		
XX	Disclosure; Page 13; 38pp; English.	
CC	The specification describes a fish serine proteinase. The proteinases are useful in mammals, for treating and preventing a disease in a human or an animal, as pain, acute inflammation, chronic inflammation, arthritis, rheumatoid arthritis, juvenile rheumatoid arthritis, systemic lupus erythematosus, osteoarthritis, fibromyalgia, acne, eczema, facial seborrheic eczema, tendonitis, rash, face or neck, foreskin infections, athlete's foot, fistula of the hands, infected topical ulcers, nail infections in newborns, wrinkle infections, keloids, boils, warts and allergic itch, hemorrhoids, wounds, wound infections, wounds from burns, fungal infections and immunological autoimmune diseases. They are also useful for removing dead or peeling skin from otherwise healthy skin, and for treating or preventing a dermatitis in which pathogenesis is caused by bacteria, virus, fungus, parasite, a protozoan or a receptor mediated binding is involved. The present sequence represents the amino acid sequence of cod trypsin isozymes.	
XX	Sequence	225 AA;
XX		

Query Match 94.7%; Score 90; DB 22; Length 225;
Best Local Similarity 90.0%; Pred. No. 1.4e-07;
Matches 18; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 IVGGYXCXHSQHQVSLNS 20
db 1 IVGGYCTXSQHQSIVSLNS 20
||||| ||||| ||||| |||||

RESULT 13
ID ABB04644 standard; Protein; 240 AA.
XX ABB04644;
DD 04-MAR-2002 (first entry)
EN Graculis japonicus trypsinogen (atry I) SEQ ID NO:1.
DE Protease; trypsinogen; sardine; Japanese anchovy, fish sauc.
XX Graculis japonicus.
SS Engraulis japonicus.
NN JP2001269173-A.
CX 02-OCT-2001.
CX 24-MAR-2000; 2000JP-0084302.
CX 24-MAR-2000; 2000JP-0084302.
XA (NISR) JAPAN TOBACCO INC.
RR MP1. 2002-078276/11.
TT N-PSDB; ABA04345.
X A new DNA sequence -
X Claim 1; Fig 7; 32pp; Japanese.

The present sequence represents a trypsinogen, which is a protease (I) isolated from Engraulis japonicus (also called Engraulis japonica or Japanese anchovy). The present invention also describes: (1) a DNA encoding (1), having a protease consisting of an amino acid sequence in which part of the amino acid residue is replaced, inserted or deleted in the amino acid sequence encoded by the above DNA and having a bio-activity substantially same as that of the expression vector in which the above DNA is recombined; (2) an expression derived protease in which a host cell transformed by the above sardine-expression vector is cultured and (1) is recovered; (4) a protease containing substantially no other protein derived from fish; and (5) the preparation of a fish sauce in a short period while inhibiting the generation of an unpleasant smell compared to a case where the sauce is prepared by the above method is not added in which at least one of the above (1)-(5) is used. The present invention relates to Clupeidae order is shellfish selected from the group consisting of as high salt concentration as 10% aqueous solution containing salts in by the above method is added to it and fermented for about 1 to 11 months. The method is used for the preparation of a fish sauce in a short period.

X Sequence 240 AA;
K

Query Match 93.2%; Score 79; DB 23; Length 240;
Best Local Similarity 75.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 5; Mismatch 5; Indels 0; Gaps 0;

/ 1 IVGGYXCXHSQHQVSLNS 20
20 IVGGYEQANSQPHVSLS 39
||||| ||||| ||||| |||||

Wed Feb 12 11:59:30 2003

us-10-036-371-1.rag

RESULT 14
 AAY33346
 ID AAY33346 standard; peptide; 25 AA.
 AC AAY33346;
 XX
 XX 29-NOV-1999 (first entry)
 DT
 DE Atlantic cod hydrolase N-terminal fragment I.
 XX
 XX Multifunctional enzyme; treatment: acne; eczema; chymotrypsin; trypsin;
 KW collagenase; elastase; exopeptidase; cell surface receptor; anti-acne;
 KW antiseborrheic; dermatological; anti-eczema; adhesion molecule; ICAM-1;
 KW CD54; ICAM-2; VCAM-1; CD4; CD8; CD28; CD29D; CD31; CD44; CD49; CD62L;
 KW CD102; GM1 ceramide; krill; atlantic cod.
 XX
 XX Gadus sp.
 OS
 XX US9598406-A.
 FN
 XX 28-SEP-1999.
 PD
 XX 08-FEB-1996; 96US-0600273.
 PF
 XX 22-NOV-1994; 94US-0388501.
 PR
 XX 06-FEB-1995; 95US-0385540.
 PR
 XX 07-JUN-1995; 95US-0486820.
 PR
 XX (PHAI-) PHAIRSON MEDICAL INC.
 PA
 XX Lindblom R, Kay J, Franklin RL, De Faire JR;
 PI
 XX WPI; 1999-561004/47.
 DR
 XX Treating acne and eczema using a krill-derived multifunctional enzyme -
 PT
 XX Disclosure; Column 21-22; 42pp; English.
 PS
 XX This invention describes a novel method for treating acne and eczema
 CC using a krill-derived multifunctional enzyme (I) which comprises or
 CC more of the activities of chymotrypsin, trypsin, collagenase, elastase
 CC or exopeptidase and is reactive with cell surface receptors such as
 CC proteins or glycoproteins. The product of the invention has activity
 CC antiseborrheic, anti-acne, dermatological and anti-eczema activity.
 CC (I) removes or inactivates cell surface receptors (proteins and
 CC glycoproteins) and adhesion molecules (ICAM-1 (i.e. CD54), ICAM-2,
 CC VCAM-1, CD4, CD8, CD28, CD29D, CD31, CD44, CD49, CD62L, CD102 and the
 CC asialo GM1 ceramide. This sequence represents the N-terminal fragment of
 CC an atlantic cod (Gadus sp.) hydrolase which is used to describe the
 CC method of the invention.
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 82.1%; Score 78; DB 20; Length 25;
 Best Local Similarity 75.0%; Pred. No. 1.6e-06;
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IVGGYKXCHSHQAOVSLNS 20
 DB 1 IVGGYCKAYSQAYQVSLNS 20
 ||||| : : : : :
 ||||| : : : : :
 RESULT 15
 AAB22965
 ID AAB22965 standard; peptide; 25 AA.
 AC AAB22965;
 XX
 XX 10-JAN-2001 (first entry)
 DT
 XX Salmon enzyme 1 multifunctional hydrolase peptide.
 DE
 XX Multifunctional hydrolase; multifunctional protein;
 KW

KW protease activity; chymotrypsin; trypsin; collagenase;
 KW elastase; wound healing; corneal ulcer; internal trauma;
 KW internal surgical wound; nerve; tendon; sheath;
 KW adhesion formation inhibition; ophthalmological; vulnery.
 XX
 OS Oncothynchus sp.
 XX
 XX WO200049991-A2.
 PN
 XX 31-AUG-2000.
 PD
 XX 29-JUN-1999; 99MO-US14751.
 PF
 XX 23-FEB-1999; 99US-0256484.
 PR
 XX (PHAI-) PHAIRSON MEDICAL INC.
 PA
 XX Franklin R, Cowling D, Hubbell JA;
 PI
 XX WPI; 2000-587120/55.
 DR
 XX Use of microgel comprising crosslinked polyanionic polymer and
 PT optionally proteolytic enzyme in treatment of corneal wounds e.g.
 PT ulcers and abrasions and internal trauma e.g. surgical wounds and for
 PT treating implants to reduce adhesions -
 PS Disclosure; Page 25; 85pp; English.
 XX
 XX The invention relates to use of a microgel comprising a crosslinked
 CC polyanionic polymer in compositions for treatment of an area affected by
 CC trauma selected from corneal wounds and internal trauma. The
 CC compositions further comprise a multifunctional protease having
 CC activities selected from two of chymotrypsin activity, trypsin activity,
 CC collagenase activity and elastase activity. In particular, the protease
 CC may be one of six isoforms of white shrimp (Penaeus vannamei)
 CC multifunctional hydrolase (AAB22947-B22952), or a multifunctional enzyme
 CC comprising one of the peptide sequences given in AAB22953-B22967.
 CC Compositions of the invention may be used to treat corneal ulcers
 CC (including infected ulcers), or abrasions or a chemical or physical
 CC insult to the cornea that are likely to give rise to a corneal ulcer.
 CC They are also used to treat internal traumas selected from an
 CC internal surgical wound, or a trauma to a membrane that covers either
 CC an internal organ or tissue or the cavity in which one or more internal
 CC organs or tissues reside, with the membrane being selected from the
 CC peritoneum, the pericardium, the epicardium and the pleura and nerve
 CC meninges. The internal trauma is to a tendon, tendon sheath, rise to
 CC or a nerve sheath, where the trauma is susceptible to adhesions and
 CC adhesions. The amount of microgel administered is sufficient to prevent
 CC or reduce formation or reformation of adhesions and the composition can
 CC also be used to treat an implant to reduce formation of adhesions.
 CC Sequences AAB22953-B22967 represent peptide sequences of
 CC multifunctional hydrolases. Multifunctional hydrolases comprising one of
 CC these sequences are claimed for use in compositions of the invention.
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 82.1%; Score 78; DB 21; Length 25;
 Best Local Similarity 75.0%; Pred. No. 1.6e-06;
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IVGGYKXCHSHQAOVSLNS 20
 DB 1 IVGGYCKAYSQAYQVSLNS 20
 ||||| : : : : :
 ||||| : : : : :
 Search completed: February 12, 2003, 10:22:20
 Job time : 14.4328 secs


```

NAME/KEY: MOD_RES
LOCATION: (71)
OTHER INFORMATION: I or unknown
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (82)
OTHER INFORMATION: N or D
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (92)
OTHER INFORMATION: K or E
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (129)
OTHER INFORMATION: A or D
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (136)
OTHER INFORMATION: V or C
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (138)
OTHER INFORMATION: N or S
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (204)
OTHER INFORMATION: H or N
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (212)
OTHER INFORMATION: M, V or C
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (324)
OTHER INFORMATION: N or S
OTHER INFORMATION: N or S
US-10-036-371-8

Query Match          97.6%; Score 1130; DB 12; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.2e-104; Indels 0; Gaps 0;
Matches 225; Conservative 0; Mismatches 0;

Qy 1 IVGGVECTHSAQHQVSLNSGTHKCGSLINXWVWSAHCYKSVLVRGLGHHIRVNEG 60
   11
Db 1 IVGGVECTHSAQHQVSLNSGTHKCGSLINXWVWSAHCYKSVLVRGLGHHIRVNEG 60
   11
Qy 61 TQKXISSSVGRHNYSYXNDIMLKLTXPATLINQYVHAVALPTECAADATMCTVSG 120
   61
Db 61 TQKXISSSVGRHNYSYXNDIMLKLTXPATLINQYVHAVALPTECAADATMCTVSG 120
   61
Qy 121 WGNWSSVWDGDKLXLPLTSLSHADANSYGPGMTTOSMFCAGYLEGGKDCQDSDGGP 180
   121
Db 121 WGNWSSVWDGDKLXLPLTSLSHADANSYGPGMTTOSMFCAGYLEGGKDCQDSDGGP 180
   121
Qy 181 WCVNGVLQGVGVSWSVGGCAERDXPGVAKVXVLGSRVDTMAY 225
   181
Db 181 WCVNGVLQGVGVSWSVGGCAERDXPGVAKVXVLGSRVDTMAY 225
   181

RESULT 2
US-10-021-368-10
Sequence 10, Application US/10021368
Patent No. US20020106367A1
GENERAL INFORMATION:
APPLICANT: Band, Vimla
TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA USA

```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-7779-134

Query Match      61.3%; Score 710; DB 10; Length 247;
Best Local Similarity 60.3%; Pred. No. 16-62;
Matches 135; Conservative 31; Mismatches 54; Indels 4; Gaps 3;

QY 1 IVGGVECTHSAQHVSINSYHKGSSLINXWYVSAACHYKSVLVRGLGHHIRVNG 60
DB 24 IVGGYICEENSVPQVNSLNSYHFCGGLISEQWVVSAGHCYKSVLVRGLGHHIRVNG 83
QY 61 TEQXISSSVXRHNPYSYNIKNDIMLKLTXPATLNOYHVALPTECAADATMCTVSG 120
DB 84 NEQFINAAKILRHFKNGRTLNDLILKLSPAVINSRVSAISLPTAPPAAGTESLISG 143
QY 121 WNTNMSYXD-GDKLQXILPILSHADCANSGFGMITQSMFCAGYLEGKDCQGDGSG 179
DB 144 WNTLSSGDDYDDELQCDAPVLSQAECASY-PGKLTNNFCVGLGKDCQGDGSG 202
QY 180 PVNCGVGLGVGVNCGVNGCAERDQPTAKVXLSGWRDPA 223
DB 203 PVVSGKLQ--GIVSWGCGAKKQKQVTKYKYNVYDKTIA 244

RESULT 4
US-09-910-071-14
; Sequence 14, Application US/09910071
; Patent No. US20020116146A1
; GENERAL INFORMATION:
; APPLICANT: Tomikawa, Mayumi
; APPLICANT: Aikawa, Seiichi
; APPLICANT: Matsuzawa, Fumiko
; TITLE OF INVENTION: Method and Apparatus for Extracting and Evaluating Mutually Similar
; TITLE OF INVENTION: Portions in One-Dimensional Sequences in Molecules and/or Three-
; FILE REFERENCE: 522.191111 Structures of Molecules
; CURRENT APPLICATION NUMBER: US/09/910, 071
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/014, 867
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 223
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: amino acid sequence of trypsin

US-09-910-071-14

Query Match      60.6%; Score 702; DB 10; Length 223;
Best Local Similarity 60.3%; Pred. No. 5.7e-62;
Matches 135; Conservative 29; Mismatches 56; Indels 4; Gaps 3;

QY 1 IVGGVECTHSAQHVSINSYHKGSSLINXWYVSAACHYKSVLVRGLGHHIRVNG 60
DB 1 IVGGYTCANTVPQVNSLNSYHFCGGLISQWVVSAGHCYKSVLVRGLGHHIRVNG 60
QY 61 TEQXISSSVXRHNPYSYNIKNDIMLKLTXPATLNOYHVALPTECAADATMCTVSG 120
DB 61 NQPTISAKSTVHPSYNSNTLNDIMLKLSASLNSRVASISLPTSCASAGTCLISG 120
QY 121 WNTNMS--YXGDKLQXILPILSHADCANSGFGMITQSMFCAGYLEGKDCQGDGSG 179
DB 121 WNTLSSGTSYDVKLCKLKPILSDSCSAGY-PQPTSNMFCAGYLEGKDCQGDGSG 179
QY 180 PVNCGVGLGVGVNCGVNGCAERDQPTAKVXLSGWRDPA 223
DB 180 PVVCSGKLQ--GIVSWGCGAKKQKQVTKYKYNVYDKTIA 221

RESULT 5
```

```
US-10-021-368-7
; Sequence 7, Application US/10021368
; Patent No. US2002010676A1
; GENERAL INFORMATION:
; APPLICANT: Band, Viola
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPILED IN: BLAST 2.0
; MEDIA: 3.5" FLOPPY disk
; OPERATING SYSTEM: PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/021,368
; FILING DATE: 12-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/201,038
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8306
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-021-368-7

Query Match      57.1%; Score 661.5; DB 12; Length 281;
Best Local Similarity 52.9%; Pred. No. 7.2e-58;
Matches 137; Conservative 27; Mismatches 56; Indels 39; Gaps 8;

QY 1 IVGGY-----ECTXHQSAHQVLSNSYHKGSSLINXWYVSAACHYKSVLVRGLGH 53
DB 24 IVGGYMTRYARTCRSSVPQVNSLNSYHFCGGLISQWVVSAGHCYKSVLVRGLGH 83
QY 54 H-----IRVNEGTQXISSSVXRHNPYSYNIKNDIMLKLTXPATL-
DB 84 NMTRYARINVLGNEQTVQSAKILRHFNYSNTLNDIMLKLSASLNSRVASISLPTSCASAGTCLISG 143
QY 100 HVALPTECAADATMCTVSGMNTMS--VXDGDGLQXILPILSHA-----DCANSY 151
DB 144 VASVPLPSSCAPAGTQCLISGNGTNSNVNPNLLQCVDAFVLPQAMTRYARDEASY 203
QY 152 GPKMTQSMFCAGYLEGKDCQGDGSGPVNCGVGLGVNCGVNGCAERDQPTAKVXLSGWRDPA 204
DB 204 -PDITNNMLCVGTEGGRGSCQGDGSGPVVCGELQ--GIVSWGTMTRYARGCAQDPA 260
QY 205 PGYAKVXLSGWRDPA 223
DB 261 PGVTKVNYVDIMQNTIA 279
```

```
RESULT 6
US-10-021-368-8
; Sequence 8, Application US/10021368
```

Wed Feb 12 11:59:46 2003

```
Patent No. US20020106367A1
GENERAL INFORMATION:
APPLICANT: Band, Vimla
TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
MOLECULES AND METHODS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/021,368
FILING DATE: 12-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/201,038
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/100002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-021-368-8
Query Match 53.5%; Score 619.5; DB 12; Length 239;
Best Local Similarity 50.2%; Pred. No. 1.1e-53;
Matches 132; Conservative 32; Mismatches 56; Indels 43; Gaps 8;
QY 1 IVGVECTXH-----SQHVLNSGVYKCGSLINXKXWVSAACHYKSVLRVLGEH 53
Db 37 IVGGHSTRIVFTCNLSLPQVLSLNSGSHFCGSLISEQWVVSAAHCYKTRIVLGEH 96
QY 54 H-----IRVNECTEQISSLVSVRHPNYSSVNXNDMLIKLTXPATL-----N 97
Db 97 NSTRIVFTIKVLGNEQFINAAKIRHPKYNKTRDLDNDMLIKLSSPAVHSTRIVTEN 156
QY 98 QYTHAVALPTECAADATMCTVSGWNTMSVYKD-GPKLOXLPIILSH-----DCA 148
Db 157 ARVSTISLTPAAGTEGLISGNTLSFGADVPDLKCLDAPVLTOAHSTRIVBECK 216
QY 149 NSYVGMITQSMFCAGYLEGKDCQDGGPVVGVNGVQVSWGY-----GCA 200
Db 217 ASY-FGKITNSMFCVLEGKDCQDGGPVVGVNGQLQ--GVWSGHSHSTRIVBGCA 273
QY 201 ERDPGVYAKVVLSGWVRTMA 223
Db 274 WKNRPGVYTKVYVYVDMIKDTIA 296
```

```
RESULT 7
US-10-021-368-9
Sequence 9, Application US/10021368
Patent No. US20020106367A1
GENERAL INFORMATION:
APPLICANT: WHYTE, DAVID
```

```
APPLICANT: Band, Vimla
TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
MOLECULES AND METHODS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/021,368
FILING DATE: 12-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/201,038
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/100002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-021-368-9
Query Match 47.5%; Score 549.5; DB 12; Length 286;
Best Local Similarity 46.0%; Pred. No. 7.8e-47;
Matches 121; Conservative 22; Mismatches 77; Indels 43; Gaps 7;
QY 1 IVGVE-----CTXHSQAVLSLNSGVYKCGSLINXKXWVSAACHYKSVLRVLGE 52
Db 25 IVGGYNTNRPVFTCBQHSVPVQSLNAGSHICGSLITDQWLSAACHYHPQLQVLGE 84
QY 53 H-----IRVNECTEQISSLVSVRHPNYSSVNXNDMLIKLTXPATL----- 96
Db 85 HNRTNRPVFTVYIEGABQFIDAAKMTLHPDYDKWTVDDNDMLIKLSPATLNTRPVB 144
QY 97 NQYTHAVALPTECAADATMCTVSGWNTMSVYKDGKLOXLPIILSHD-----CA 148
Db 145 NSKVSTIPLPQYCTAGTECLVSGWVLPKGFSPSVLQCLDAPVLSRNTNRPVBVCH 204
QY 149 NSYVGMITQSMFCAGYLEGKDCQDGGPVVGVNGVQVSWG-----YCCA 200
Db 205 KAY-PKQITNNMFCVLEGKDCQDGGPVVGVNGEVQ--GIVSGDRNTRPVBGCA 261
QY 201 ERDPGVYAKVVLSGWVRTMA 223
Db 262 LEGPGVYTKVYVYVDMIKDTIA 284
```

```
RESULT 8
US-09-888-615-90
Sequence 90, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
```

```

: APPLICANT: CHENEPEL, SEAN
: APPLICANT: CHARYOCZAK, JEN
: APPLICANT: MANNING, GREGORY
: APPLICANT: SUDARSANAM, SUCHA
: TITLE OF INVENTION: NOVEL PROTEASES
: CURRENT APPLICATION NUMBER: US/09/888,615
: CURRENT FILING DATE: 2001-06-26
: PRIOR APPLICATION NUMBER: 60/214,047
: PRIOR FILING DATE: 2000-06-26
: NUMBER OF SEQ ID NOS: 150
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 90
: LENGTH: 240
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-888-615-90

Query Match 42-78; Score 494; DB 10; Length 320;
Best Local Similarity 44-78; Pseq No. 2,66-44;
Matches 105; Conservative 27; Mismatches 857; Indels 16; Gaps 5;

QY 1 IWGVGECTXHSQHQVSL-NSGVHKGCGSLINXWVSNAAHYCVLRLVGEHHRVNE 59
Db 86 LLEEGDCAPHSPQWVALVEGRFNCASLSPHWLSAAHCSQSFHVLGEHNRKED 145
QY 60 GTEGXISSSVSRHPNTSSYNXNDIMLILKLTPTALNQYVHAVALPTCAADATMCTCS 119
Db 146 GPEQLATTSRVIPHRFYEARSHRNDIMLRLVQAPRLNPQVRPVLTPCRHPGEACVNS 205
QY 120 GW-----GNTMSSVXDGDQKLOKLPILSHADCANISYVGSGMITQSMFCAGYLEG 168
Db 206 GNGLVSHNEFGTASPRQSVSLPTLTHANISIIISDTSCKASY-EGRLNTMTWCAQAGSR 264
QY 169 GKSGCGDSGPNVQVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 222
Db 265 GAESCEGDSGGLVCGGLQ--GVSGNDVPCDNTKGVGVGVGVGVGVGVGVGVGVGVGVGV 317

RESULT 9
US-10-028-072-506
: Sequence 506, Application US/10028072
: Publication No. US20030004311A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Fennoyers, Luc
: APPLICANT: Filvaroi, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Goddard, Jeffrey E.
: APPLICANT: Goddard, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tamas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang
: TITLE OF INVENTION:
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/10/028,072
: CURRENT FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: 60/049911
: PRIOR FILING DATE: 1997-06-18
: PRIOR APPLICATION NUMBER: 60/056974
: PRIOR FILING DATE: 1997-06-18
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059115
: PRIOR FILING DATE: 1997-09-17

```

1 PRIOR FILING DATE: 1998-01-23
2 PRIOR APPLICATION NUMBER: 60/073612
3 PRIOR FILING DATE: 1998-02-04
4 PRIOR APPLICATION NUMBER: 60/074086
5 PRIOR FILING DATE: 1998-02-09
6 PRIOR APPLICATION NUMBER: 60/074092
7 PRIOR FILING DATE: 1998-02-09
8 PRIOR APPLICATION NUMBER: 60/077791
9 PRIOR FILING DATE: 1998-03-20
10 PRIOR APPLICATION NUMBER: 60/078910
11 PRIOR FILING DATE: 1998-03-20
12 PRIOR APPLICATION NUMBER: 60/079294
13 PRIOR FILING DATE: 1998-03-25
14 PRIOR APPLICATION NUMBER: 60/079663
15 PRIOR FILING DATE: 1998-02-27
16 PRIOR APPLICATION NUMBER: 60/079728
17 PRIOR FILING DATE: 1998-03-27
18 PRIOR APPLICATION NUMBER: 60/080165
19 PRIOR FILING DATE: 1998-03-31
20 PRIOR APPLICATION NUMBER: 60/081203
21 PRIOR FILING DATE: 1998-04-09
22 PRIOR APPLICATION NUMBER: 60/081229
23 PRIOR FILING DATE: 1998-04-09
24 PRIOR APPLICATION NUMBER: 60/081695
25 PRIOR FILING DATE: 1998-04-14
26 PRIOR APPLICATION NUMBER: 60/081817
27 PRIOR FILING DATE: 1998-04-15
28 PRIOR APPLICATION NUMBER: 60/081818
29 PRIOR FILING DATE: 1998-04-15
30 PRIOR APPLICATION NUMBER: 60/082999
31 PRIOR FILING DATE: 1998-04-24
32 PRIOR APPLICATION NUMBER: 60/083322
33 PRIOR FILING DATE: 1998-04-28
34 PRIOR APPLICATION NUMBER: 60/083545
35 PRIOR FILING DATE: 1998-04-29
36 PRIOR APPLICATION NUMBER: 60/084600
37 PRIOR FILING DATE: 1998-05-07
38 PRIOR APPLICATION NUMBER: 60/084627
39 PRIOR FILING DATE: 1998-05-07
40 PRIOR APPLICATION NUMBER: 60/084637
41 PRIOR FILING DATE: 1998-05-07
42 PRIOR APPLICATION NUMBER: 60/085149
43 PRIOR FILING DATE: 1998-05-12
44 PRIOR APPLICATION NUMBER: 60/085323
45 PRIOR FILING DATE: 1998-05-11
46 PRIOR APPLICATION NUMBER: 60/085338
47 PRIOR FILING DATE: 1998-05-11
48 PRIOR APPLICATION NUMBER: 60/085339
49 PRIOR FILING DATE: 1998-05-11
50 PRIOR APPLICATION NUMBER: 60/085579
51 PRIOR FILING DATE: 1998-05-15
52 PRIOR APPLICATION NUMBER: 60/085697
53 PRIOR FILING DATE: 1998-05-15
54 PRIOR APPLICATION NUMBER: 60/085704
55 PRIOR FILING DATE: 1998-05-15
56 PRIOR APPLICATION NUMBER: 60/086414
57 PRIOR FILING DATE: 1998-05-22
58 PRIOR APPLICATION NUMBER: 60/086430
59 PRIOR FILING DATE: 1998-05-22
60 PRIOR APPLICATION NUMBER: 60/087106
61 PRIOR FILING DATE: 1998-05-28
62 PRIOR APPLICATION NUMBER: 60/088026
63 PRIOR FILING DATE: 1998-06-04
64 PRIOR APPLICATION NUMBER: 60/088730
65 PRIOR FILING DATE: 1998-06-10
66 PRIOR APPLICATION NUMBER: 60/088741
67 PRIOR FILING DATE: 1998-06-10
68 PRIOR APPLICATION NUMBER: 60/088810
69 PRIOR FILING DATE: 1998-06-10
70 PRIOR APPLICATION NUMBER: 60/088959
71 PRIOR FILING DATE: 1998-06-11
72 PRIOR APPLICATION NUMBER: 60/089532
73 PRIOR FILING DATE: 1998-06-17

1 PRIOR APPLICATION NUMBER: 60/089599
2 PRIOR FILING DATE: 1998-06-17
3 PRIOR APPLICATION NUMBER: 60/089907
4 PRIOR FILING DATE: 1998-06-18
5 PRIOR APPLICATION NUMBER: 60/089947
6 PRIOR FILING DATE: 1998-06-19
7 PRIOR APPLICATION NUMBER: 60/090349
8 PRIOR FILING DATE: 1998-06-23
9 PRIOR APPLICATION NUMBER: 60/090429
10 PRIOR FILING DATE: 1998-06-24
11 PRIOR APPLICATION NUMBER: 60/090445
12 PRIOR FILING DATE: 1998-06-24
13 PRIOR APPLICATION NUMBER: 60/090538
14 PRIOR FILING DATE: 1998-06-24
15 PRIOR APPLICATION NUMBER: 60/090863
16 PRIOR FILING DATE: 1998-06-26
17 PRIOR APPLICATION NUMBER: 60/091360
18 PRIOR FILING DATE: 1998-07-01
19 PRIOR APPLICATION NUMBER: 60/091519
20 PRIOR FILING DATE: 1998-07-02
21 PRIOR APPLICATION NUMBER: 60/091982
22 PRIOR FILING DATE: 1998-07-07

Query Match 38 94; Score 451, DB 9; Length 250;
Best Local Similarity 41 04; Pred. No. 3,3e-37;
Matches 94; Conservative 38; Mismatches 87; Indels 10; Gaps 6;

Qy 1 IIVGVECTYHSQAHQVSL-NSGYHXGSGSLNXXVWVAHCHYKSVLRVLRGHEHVRNE 59
Db 22 IIKGFECKPSHQWQALFEKTRLLCGATLTAPEWLLTAHCLKPRYIVHLGHNLOKEE 81
Qy 60 GTEQXISSSVXRHNYS---SYNIXNDIMLIXTPATLNQVHVAALPTECAADATM 115
Db 82 GCQTRTATSFPHGPFNNSLPNKDHNDIMLVQWASPVISITWAVRPTLSSRCVTAGTS 141
Qy 116 CTVSGMGNTWS-SVADGDKLQXLPLSLHADCANSGYFGMITQSCFAGYLGKQDSQ 174
Db 142 CLISWGSGTSSPOLRPHLTRCANITIIHQKCNAY-PGNITDTWVCASVQGGKQDSQ 200
Qy 175 GDSGGPVVGVNLQGVVWSWYG-CAERDXPGVYAKVXVLSGNVDMT 222
Db 201 GDSGGFLVCNQLQ--GIISWGQDFCAITRKPGVYTKVCKYVDMIGETM 247

RESULT 10
US-10-121-049-506
Sequence 506, Application US/10121049
Publication No. US2003002239A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Deenoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, V. Forrest
APPLICANT: Steward, David
APPLICANT: Tumas, Robert
APPLICANT: Wood, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P333081C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File wrapper or Palm
NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 506

; LENGTH: 250

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-121-049-506

Query Match 38.9%; Score 451; DB 9; Length 250;

Best Local Similarity 41.0%; Pred. No. 3.3e-37;

Matches 94; Conservative 38; Mismatches 87; Indels 10; Gaps 6;

QY 1 IVGGVECTHSQAHQVSL-NSGVHXCGLINXWVWSAHCYKSVLRVLGEHHRVNE 59

Db 22 IIKGFCKPHSQPQAALFETRLLCATLIAPRMLLTAAHCLKPRYIVHLQGNLQKEE 81

QY 60 GTEQXISSSVXRHPNYS---SYNIKNDIMLKLTPTALNOYVAVALPTCAADATM 115

Db 82 GCEQTRTATESFPHPGFNNSLPNKDHNDIMLVKMASFVSITWAVRPLTSSRCVTAGTS 141

QY 116 CTVSGNGNTMS-SVXGDKLQXLLPILSHADCANSYGPQMITSQMFCAGYLEGKDSQ 174

Db 142 CLISMGSTSSPQLRPHLTLCANITIIHQKCNAY-PGNITDTMVCASVQEGKDSQ 200

QY 175 GDSGPPVNCVGLGVGVWSGVG-CAERDXPGVYAKYVLSGWRDTM 222

Db 201 GDSGGLVNCQSLQ--GIISWGDDPCAITRKPGVYTKVKYVDMIQETM 247

RESULT 12

US-10-140-470-506

; Sequence 506, Application US/10123904

; Publication No. US2003002238A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: ACTUS ENCODING THE SAME

; CURRENT APPLICATION NUMBER: US/10/123,904

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 506

; LENGTH: 250

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-123-904-506

Query Match 38.9%; Score 451; DB 9; Length 250;

Best Local Similarity 41.0%; Pred. No. 3.3e-37;

Matches 94; Conservative 38; Mismatches 87; Indels 10; Gaps 6;

QY 1 IVGGVECTHSQAHQVSL-NSGVHXCGLINXWVWSAHCYKSVLRVLGEHHRVNE 59

Db 22 IIKGFCKPHSQPQAALFETRLLCATLIAPRMLLTAAHCLKPRYIVHLQGNLQKEE 81

QY 60 GTEQXISSSVXRHPNYS---SYNIKNDIMLKLTPTALNOYVAVALPTCAADATM 115

Db 82 GCEQTRTATESFPHPGFNNSLPNKDHNDIMLVKMASFVSITWAVRPLTSSRCVTAGTS 141

QY 116 CTVSGNGNTMS-SVXGDKLQXLLPILSHADCANSYGPQMITSQMFCAGYLEGKDSQ 174

Db 142 CLISMGSTSSPQLRPHLTLCANITIIHQKCNAY-PGNITDTMVCASVQEGKDSQ 200

QY 175 GDSGPPVNCVGLGVGVWSGVG-CAERDXPGVYAKYVLSGWRDTM 222

Db 201 GDSGGLVNCQSLQ--GIISWGDDPCAITRKPGVYTKVKYVDMIQETM 247

RESULT 12

US-10-140-470-506

; Sequence 506, Application US/10140470

; Publication No. US2003002231A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Matanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: ACTUS ENCODING THE SAME

; CURRENT APPLICATION NUMBER: US/10/140,470

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 506

; LENGTH: 250

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-140-470-506

Query Match 38.9%; Score 451; DB 9; Length 250;

Best Local Similarity 41.0%; Pred. No. 3.3e-37;

Matches 94; Conservative 38; Mismatches 87; Indels 10; Gaps 6;

QY 1 IVGGVECTHSQAHQVSL-NSGVHXCGLINXWVWSAHCYKSVLRVLGEHHRVNE 59

Db 22 IIKGFCKPHSQPQAALFETRLLCATLIAPRMLLTAAHCLKPRYIVHLQGNLQKEE 81

QY 60 GTEQXISSSVXRHPNYS---SYNIKNDIMLKLTPTALNOYVAVALPTCAADATM 115

Db 82 GCEQTRTATESFPHPGFNNSLPNKDHNDIMLVKMASFVSITWAVRPLTSSRCVTAGTS 141

QY 116 CTVSGNGNTMS-SVXGDKLQXLLPILSHADCANSYGPQMITSQMFCAGYLEGKDSQ 174

Db 142 CLISMGSTSSPQLRPHLTLCANITIIHQKCNAY-PGNITDTMVCASVQEGKDSQ 200

QY 175 GDSGPPVNCVGLGVGVWSGVG-CAERDXPGVYAKYVLSGWRDTM 222

Db 201 GDSGGLVNCQSLQ--GIISWGDDPCAITRKPGVYTKVKYVDMIQETM 247

RESULT 13

US-10-175-746-506

; Sequence 506, Application US/10175746

; Publication No. US20030027270A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

Db 22 !:KGFEKPHSOPMAALFEKTELLCGATLIAPRWLLTAHCLKPRYIVHLGQHNLOKEE 81
QY 60 GTEQXISSSVXRHPNYS---SYNIXNDIMLIKLTXPATLNQYVHVALEFTECAADATM 115
Db 82 GCEQFTATIESPFHGFNNLSLFPKQHRNDIMLVQMASFVSITWAVRELLSSRCVTAGTS 141
QY 116 CTVSGMGNTWS-SVXDGDYLOXLXLPILSHADGANSYGPQMITSQNECAGYLEGKDSQ 174
Db 142 CLISNGSTSSPOLRPLHLRCAN:IIIEHOKCENAY-PGNITDTWNCASVOEGKDSQ 200
QY 175 GDSGQFVYVNCVLQGVVWSMGY-CAERDXPGYAKTVLSSAVRDTM 222
Db 201 GDSGQFLVCNQLQ--GIISWGQDPCAITRKEGVYTKVKYVDMIQETM 247

Search completed: February 12, 2003, 10:31:08
Job time : 35.2537 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:16:06 ; Search time 159.179 Seconds
(without alignments)
291.248 Million cell updates/sec

Title: US-10-036-371-8

Perfect score: 1158

Sequence: 1 IVGGVCTXHSQHQVSLNS.....GVYAKVKVLSQHWRTWAXY 225

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_mus.*
9: sp_organelle.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvitus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932.5	80.5	219	13	Q91036
2	915.5	79.1	242	13	Q9W707
3	910.5	78.6	242	13	Q93266
4	898.5	77.6	237	13	Q91515
5	897.5	77.5	242	13	Q92099
6	887.5	76.6	238	13	Q9W7Q6
7	886	76.5	240	13	Q987G9
8	882	76.2	241	13	Q98TG9
9	844.5	72.9	244	13	Q9QW33
10	740.5	63.9	244	13	Q42159
11	737.5	63.7	247	13	Q42608
12	736.5	63.6	245	13	Q42608
13	732	63.6	247	13	Q42158
14	732	63.6	246	11	Q9QK99
15	735	63.3	246	11	Q9QK77
16	733.5	63.3	178	13	Q93594

17	729	63.0	246	11	Q92LR9
18	715	61.7	247	11	Q92LR9
19	711	61.4	247	11	Q92LR9
20	659.5	57.0	247	11	Q92LR9
21	640	55.3	247	13	Q9W7Q5
22	587	50.7	250	13	Q93265
23	581.5	50.2	249	13	Q9MEK0
24	563.5	48.7	675	13	Q9MEJ8
25	560.5	48.4	249	13	Q92046
26	548.5	47.4	344	13	Q9MEJ9
27	498	43.0	264	5	Q62561
28	493	42.6	255	4	Q96R00
29	488	42.1	263	5	Q62562
30	474	40.9	263	5	Q97Y16
31	474	40.9	266	5	Q27761
32	471	40.7	262	5	Q9BK47
33	471	40.7	274	5	Q9K413
34	467.5	40.4	274	5	Q9K413
35	467.5	40.4	233	11	Q9D140
36	461.5	39.9	235	11	Q63274
37	458.5	39.6	274	5	Q17086
38	450	38.9	267	5	Q9VLF5
39	442.5	38.2	234	11	Q9CV76
40	442	38.2	249	11	Q9QVY4
41	442	38.2	276	11	Q9QVY3
42	439	37.9	146	13	Q9DDE1
43	438.5	37.9	239	11	Q63275
44	437	37.7	268	5	Q46151
45	436.5	37.7	264	5	Q02569

ALIGNMENTS

RESULT 1

ID	Q91036	PRELIMINARY;	PRT;	219 AA.
AC	Q91036;			
DT	01-NOV-1996	(TRENBLrel. 01, Created)		
DT	01-NOV-1996	(TRENBLrel. 01, Last sequence update)		
DT	01-JUN-2002	(TRENBLrel. 21, Last annotation update)		
DE	Trypsinogen 1 (Fragment).			
OS	Gadus morhua (Atlantic cod).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	13__taxid8049;			
RN	SEQUENCE FROM N.A.			
RA	Org T.L. Armstrong R.P. McNamara P. Buckley L.J.;			
RL	Submitted (CAN-1996) to the EMBL/GenBank/DBS databases.			
CC	-1- SIMILARITY: BELONGS TO TRYPSIN FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY			
DR	EMBL; U47819; AAB02196.1; ..			
DR	HSSP; P00763; LDPO.			
DR	MEROPS; S01.125; ..			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00200; Tryp_Spc; 1.			
DR	PROSITE; PS00240; TRYP SIN DOM; 1.			
DR	PROSITE; PS00134; TRYP SIN HIS; UNKNOWN 1.			
DR	PROSITE; PS00135; TRYP SIN SER; 1.			
KW	Hydrolase; Serine protease.			
FT	NON_TER			
SQ	SEQUENCE 219 AA; 23525 MW; C96964EB49CEDIDA CRC64;			

Query Match 80.5%; Score 932.5; DB 13; Length 219;
Best Local Similarity 87.6%; Pred. No. 1.e-90;
Matches 176; Conservative 5; Mismatches 17; Indels 3; Gaps 2;

QY 1 IVGGVCTXHSQHQVSLNSGYHXCGLS.LNXWVVAHCVKSLVRLVRLGHHHLPWNEG 60

```
Db 20 IVGGVCTKHSHQHQVSLNSGVHFCGSLVSKDMVSAACHYKSRVRLGHEHHRVNEG 79
Qy 61 TEQXISSSVXKRPNSYSSNINDIMLIKTPATLNQYVAVALPTCAADATMCTVSG 120
Db 80 TEGYISSSVIRHPNSYSSNINDIMLIKSPATLNQYVAVALPTCAADATMCTVSG 139
Qy 121 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 180
Db 140 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 198
Qy 181 VVNCNGVLQGVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 201
Db 199 VVNCNGVLQ--GVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 217

RESULT 2
Q9W707 PRELIMINARY; PRT; 242 AA.
AC Q9W707;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen 1.
OS Paralicthys olivaceus (Plounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RA "Japanese flounder mRNA for trypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB028750; BAA82362.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR HYDROLase; Serine protease; Signal.
KW HYDROLase; Serine protease; Signal.
FT CHAIN 1..242 TRYPSINogen 2.
SQ SEQUENCE 242 AA; 26546 MW; 6D4722C80BC194A2 CRC64;

Query Match 79.1%; Score 915.5; DB 13; Length 242;
Best Local Similarity 76.4%; Pred. No. 7.9e-89;
Matches 17; Conservative 11; Mismatches 39; Indels 3; Gaps 2;

Qy 1 IVGGVCTKHSHQHQVSLNSGVHFCGSLINXWVWSAAHCYKSVRLGHEHHRVNEG 60
Db 21 IVGGVCTKHSHQHQVSLNSGVHFCGSLVNNWVWSAAHCYKSVRLGHEHHRVNEG 80
Qy 61 TEQXISSSVXKRPNSYSSNINDIMLIKTPATLNQYVAVALPTCAADATMCTVSG 120
Db 81 TEQFISSSVIRHPNSYSSNINDIMLIKSPATLNQYVAVALPTCAADATMCTVSG 140
Qy 121 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 180
Db 141 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 199
Qy 181 VVNCNGVLQGVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 225
Db 200 VVNCNGELQ--GVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 242

. . .
```

```
RESULT 3
Q93266 PRELIMINARY; PRT; 242 AA.
AC Q93266;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen 2 precursor (SC 3.4.21.4).
GN Trp2.
OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.
OX NCBI_TaxID=8265;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE, PYLORIC CAECA, AND PANCREAS;
RA Douglas S.E., Gallant J.W.;
RA "Isolation of cDNAs for Trypsinogen from the Winter Flounder,
RT Pleuronectes americanus.";
RL J. Mar. Biotechnol. 0:0-0(1998).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF012463; AAC32752.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR HYDROLase; Serine protease; Signal.
KW HYDROLase; Serine protease; Signal.
FT CHAIN 1..242 TRYPSINogen 2.
SQ SEQUENCE 242 AA; 26180 MW; 08D2A834F8289080 CRC64;

Query Match 78.6%; Score 910.5; DB 13; Length 242;
Best Local Similarity 75.6%; Pred. No. 2.7e-88;
Matches 170; Conservative 15; Mismatches 37; Indels 3; Gaps 2;

Qy 1 IVGGVCTKHSHQHQVSLNSGVHFCGSLINXWVWSAAHCYKSVRLGHEHHRVNEG 60
Db 21 IVGGVCTKHSHQHQVSLNSGVHFCGSLVNNWVWSAAHCYKSVRLGHEHHRVNEG 80
Qy 61 TEQXISSSVXKRPNSYSSNINDIMLIKTPATLNQYVAVALPTCAADATMCTVSG 120
Db 81 TEQFISSSVIRHPNSYSSNINDIMLIKSPATLNQYVAVALPTCAADATMCTVSG 140
Qy 121 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 180
Db 141 WGMTSSVXDGDKLQKLXLPILSHADCANSGPGMTQSMFCAGYLEGKDCSCQDSSGP 199
Qy 181 VVNCNGVLQGVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 225
Db 200 VICNGELQ--GVVSWGCGAERDPGVYAKYKXVLSGWRDTHMAXY 242

RESULT 4
Q91515 PRELIMINARY; PRT; 237 AA.
AC Q91515;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Trypsinogen (fragment).
OS Pugu rubripes (Lapage).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
```

[illegible]

```

QY 1 IVGGYECTXHSQHQVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLGHEHVRNEG 60
Db 18 IVGGYECTPYSPHQVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLRGHEHVRNEG 77
QY 61 TEQXISSSVXKRPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 120
Db 78 TEQFISSSVIRHPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 137
QY 121 WGMTSSVXGDKLQXLPILSHADCANSGPMITQSMFCAGYLEGKDCSCQDSGGP 180
Db 138 WGMTSSV-TDSSRLQCLDLPILSERDCENSY-PGMITNMFACGLEGKDCSCQDSGGP 195
QY 181 VVNCVGLQGVVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 225
Db 196 VVNCVGLQ--GVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 238

```

```

RESULT 7
QY8THO PRELIMINARY; PRT; 240 AA.
AC QY8THO; 2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DR 01-JUN-2001 (TREMBlrel. 21, Last annotation update)
DE Trypsinogen II.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OC Engraulis.
NCBI_TaxID=42892;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
EMBL; AB041929; BAB40329.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.125; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 240 AA; 46086 MW; 299B119BFF071464 CRC64;

```

Query Match 76.5%; Score 866; DB 13; Length 240;
 Best Local Similarity 74.9%; Pred. No. 1e-85;
 Matches 167; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

```

QY 1 IVGGYECTXHSQHQVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLGHEHVRNEG 60
Db 20 IVGGYEQCAHSQPHVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLRGHEHVRNEG 79
QY 61 TEQXISSSVXKRPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 120
Db 80 TEQFISSSVIRHPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 139
QY 121 WGMTSSVXGDKLQXLPILSHADCANSGPMITQSMFCAGYLEGKDCSCQDSGGP 180
Db 140 WGMTSSV-SGDKLQCLQIPILSDRDCNSY-PGMITNMFACGLEGKDCSCQDSGGP 197
QY 181 VVNCVGLQGVVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 223
Db 196 VVNCVGLQ--GVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 238

```

```

RESULT 8
QY8THO PRELIMINARY; PRT; 241 AA.
AC QY8THO; 2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DR 01-JUN-2001 (TREMBlrel. 21, Last annotation update)
DE Trypsinogen II.
OS Engraulis japonicus (Japanese anchovy).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Engraulidae;
OC Engraulis.
NCBI_TaxID=42892;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PYLORIC CAECA;
RA Watabe S., Ahsan M.N., Funabara D.;
RT "Anchovy trypsinogen mRNA.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
EMBL; AB041930; BAB40330.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.259; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 241 AA; 46282 MW; FE362D33CAEEB2F6 CRC64;

```

Query Match 76.2%; Score 882; DB 13; Length 241;
 Best Local Similarity 73.8%; Pred. No. 2.8e-85;
 Matches 166; Conservative 15; Mismatches 40; Indels 4; Gaps 3;

```

QY 1 IVGGYECTXHSQHQVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLGHEHVRNEG 60
Db 20 IVGGYEQCPYSPHQVSLNSGYHXCGLINXWVWSAAHCYKSVLRVLRGHEHVRNEG 79
QY 61 TEQXISSSVXKRPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 120
Db 80 TEQFISSSVIRHPNSVSYNDIMLIKLPATLNQVVAVALPTECAADATMCTVSG 139
QY 121 WGMTSSVXGDKLQXLPILSHADCANSGPMITQSMFCAGYLEGKDCSCQDSGGP 180
Db 140 WGMTSSV-SGDKLQCLQIPILSDRDCNSY-PGMITNMFACGLEGKDCSCQDSGGP 197
QY 181 VVNCVGLQGVVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 225
Db 196 VVNCVGLQ--GVWSVNGYCAERDHPGVYAKVYKVLGSGWVRDTHMY 240

```

```

RC TISSUE-PANCREAS;
RA Kurokawa T., Suzuki T., Ohta H., Kagawa H., Tanaka H., Unuma T.;
RA Expression of pancreatic enzyme genes during the early larval stage
PT of Japanese eel, Anguilla japonica.";
RL Subj: AB070720, BNC85634.1;
RL DB EMBL;
RW Hydrolyase.
SW SEQUENCE. 244 AA; 26317 MW; 0EB3B68E8706D52D CRC64;

Query Match 72.9%; Score 84.5; DB 13; Length 244;
Best Local Similarity 70.7%; Pred. No. 2.6e-81;
Matches 159; Conservative 21; Mismatches 40; Indels 5; Gaps 3;

QY 1 IVGGVECTYHSQAQVSLNSGYHFCGSLNXXVWSAAHCYKSVLR--VRLGEGHHRVN 58
DB 21 IVGGVECTYHSQAQVSLNSGYHFCGSLNXXVWSAAHCYKSVLR--VRLGEGHHRVN 58
DB 21 IVGGVECTYHSQAQVSLNSGYHFCGSLNXXVWSAAHCYKSVLR--VRLGEGHHRVN 80
QY 59 ECTEQXISSSVYHNPYSSVNXNDIMLKLTYPATLNQYHVAVALPTECAADATMCTV 118
DB 81 ECTEQXISSSVYHNPYSSVNXNDIMLKLTYPATLNQYHVAVALPTECAADATMCTV 118
DB 81 ECTEQXISSSVYHNPYSSVNXNDIMLKLTYPATLNQYHVAVALPTECAADATMCTV 140
QY 119 SGMGNTSSVYDQKLOKXLPILSHADCANSGYGMITQSMFCAGYLEGGKDCQGDGSG 178
DB 141 SGMGNTSSVYDQKLOKXLPILSHADCANSGYGMITQSMFCAGYLEGGKDCQGDGSG 178
DB 141 SGMGNTSSVYDQKLOKXLPILSHADCANSGYGMITQSMFCAGYLEGGKDCQGDGSG 199
QY 179 GPVVCNGVQLGVGVWSGVCGRDXPGYAKVXVLSQWRDTMA 223
DB 200 GPVVCNGVQLGVGVWSGVCGRDXPGYAKVXVLSQWRDTMA 242

RESULT 11
Q42608 PRELIMINARY; PRT; 247 AA.
ID Q42608;
AC Q42608;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Trypsinogen A3 precursor.
GN TRYPA3.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RA [1]
RA SEQUENCE FROM N.A.
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=ANTERIOR INTESTINE;
RA Roach J.C.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF011899; AAB69655.1; -.
DR EMBL; AF011352; AAB65411.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.128; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00200; TYPD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 247 TRYPSIN A3.
SQ SEQUENCE 247 AA; 26295 MW; BECD3069A071DCB CRC64;

Query Match 63.7%; Score 737.5; DB 13; Length 247;
Best Local Similarity 62.2%; Pred. No. 5.6e-70;
Matches 140; Conservative 21; Mismatches 59; Indels 5; Gaps 3;

QY 1 IVGGVECTYHSQAQVSLNSGYHFCGSLNXXVWSAAHCYKSVLR--VRLGEGHHRVN 58
DB 24 IVGGVECTYHSQAQVSLNSGYHFCGSLNXXVWSAAHCYKSVLR--VRLGEGHHRVN 83
QY 59 ECTEQXISSSVYHNPYSSVNXNDIMLKLTYPATLNQYHVAVALPTECAADATMCTV 118
DB 84 ECTEQXISSSVYHNPYSSVNXNDIMLKLTYPATLNQYHVAVALPTECAADATMCTV 143
QY 119 SGMGNTSSVYDQKLOKXLPILSHADCANSGYGMITQSMFCAGYLEGGKDCQGDGSG 178
DB 144 SGMGNTSSVYDQKLOKXLPILSHADCANSGYGMITQSMFCAGYLEGGKDCQGDGSG 202
QY 179 GPVVCNGVQLGVGVWSGVCGRDXPGYAKVXVLSQWRDTMA 223

```

OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
† † † † †

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Katschmann M., Gaasterland H.A., Ashburner M., Batalov S., Casavant T.,
 RA Saito K., Matsuda K., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Sakai K., Matsuda K., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Schirali L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto Y., Storch K.-F.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, AND SPLEEN;
 RC MEDLINE=21095660; PubMed=11217851;
 RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kado K., Matsuda K., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Katschmann M., Gaasterland H.A., Ashburner M., Batalov S., Casavant T.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirali L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto Y., Storch K.-F.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RC MEDLINE=21103195; PubMed=11160223;
 RX Chen F., Rowen L., Hood L., Rothenberg E.V.,
 RT "Differential transcriptional regulation of individual TCR Vbeta
 RT segments before gene rearrangement.";
 RL J. Immunol. 166:1771-1780(2001)
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR ENBL; AB009661; BAA85187.1; -;
 DR ENBL; AB017031; BAA74760.1; -;
 DR ENBL; AK007843; BAB25300.1; -;
 DR ENBL; AE000664; AAB69056.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.058; -;
 DR WGI; 1913350; -;
 DR WGI; 1913350; -;
 DR InterPro; IP0001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00200; Tryp_Spc. 1.
 DR PROSITE; PS00240; TRYPSIN_DOM. 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 246 AA; 26277 MW; 915C92353EC42809 CRC64;
 Query Match 63.6%; Score 736; DB 11; Length 246;
 Best Local Similarity 62.1%; Pred. No. 8e-70;
 Matches 139; Conservative 26; Mismatches 55; Indels 4; Gaps 3;
 QY 1 VGGVETCHSQAHQVNSGTHKCGSLINXWVSAGHCYSVLVRGHERIVNEG 60
 DB 24 IVGGITCRNSIPVQVNSGTHKCGSLINXWVSAGHCYSVLVRGHERIVNEG 83
 QY 61 TEQXISSSVYHNPVSSVYNDIMLKLXTPATINQVHVALPTCCADATCTVSG 120
 DB 84 NEQFVNSAKIKHFNRSNTLNDIMLKASPVILNARVATVALPSCAPACTQCLISG 143
 QY 121 WNTWMS-SVKQDDKQLXLPILSHADCNISGPGMITOSNFCAGYLEGKDSQGS 179
 DB 144 WNTLSFGVNPDLQCLDAPLPQADCEASY-PGKITNNMIVGFLGKDSQGS 202
 QY 180 PWCNVGVLGVNSGYYGCAERDPGVYKATXVLSGWRDTMA 223
 DB 203 PWCNVGQLQ--GIWVSNGYGLKALNPGVYTKVNCVYVDIQTDTA 244
 RESULT 15
 Q9R077 PRELIMINARY; PRT; 246 AA.
 AC Q9R077
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

Qy 1 IVGGYECTXHSQAHQVSLNSGTHKCGSLINXXVVSNAHCYKSVLRVLGEHHIRVNEG 60
Db 24 IVGGYTCRENSVPYQVLSNGHFCGSLINDOWVVSAACTSR:QVRLGEHINIVLEG 83
Qy 61 TEQXISSSSVRHPYSSYINXNDIMIKITYPATINOVYHVALEPCADATMCTVSG 120
Db 84 NEGVSNAKIKRPNFNRITLNDIMLIKASPTILNARVATVALPSSCAPAGTQCLISG 143
Qy 121 NGHTAS-SVXDGDKLOXILPILSHADCANSGPMITOSMFCAGYLEGGKDCOGDSSG 179
Db 144 WGNILSGVNRPDLLQCLDAPILPONCEASY-PGKITNNMICVGFLEGGKDCOGDSSG 202
Qy 180 PVVCGVUGVGVVNSGVYGCABEDXPVYAKVXVLGWRDTMA 223
Db 203 PVVCGQLQ--GIVSWGYGALKDNPGVYTKVYVDMIONTIA 244

Search completed: February 12, 2003, 10:27:30
Job time : 160.179 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:19:40 : Search time 49.3582 Seconds
(with alignments)
136.898 Million cell updates/sec

Title: US-10-036-371-8

Perfect score: 1158

Sequence: 1 IVGGYECTVHSHQVSLNS.....GVVAKVAVLSQWRDINAXY 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/prodata//iaa/5A_COMB.pep.*

2: /cgn2_6/prodata//iaa/5B_COMB.pep.*

3: /cgn2_6/prodata//iaa/6A_COMB.pep.*

4: /cgn2_6/prodata//iaa/6B_COMB.pep.*

5: /cgn2_6/prodata//iaa/6C_COMB.pep.*

6: /cgn2_6/prodata//iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	823.5	71.1	271	1	US-08-467-155A-10
2	823.5	71.1	271	2	US-08-628-198-10
3	823.5	71.1	271	3	US-09-201-038-10
4	823.5	71.1	271	4	PCT-US96-07343-10
5	755	65.2	271	5	US-08-978-4048-44
6	710	61.3	226	2	US-08-944-483-34
7	707	61.1	247	2	US-08-766-267A-2
8	702	60.6	224	2	US-08-944-483-12
9	702	60.6	224	4	US-09-236-219-13
10	702	60.6	225	2	US-09-027-337-5
11	702	60.6	225	4	US-09-644-600-5
12	702	60.5	223	1	US-08-278-091-9
13	701	60.5	223	1	US-08-483-859-9
14	701	60.5	223	1	US-08-472-173-9
15	701	60.5	223	2	US-08-482-816-9
16	701	60.5	223	2	US-08-482-816-9
17	701	60.5	223	2	US-08-801-499-9
18	701	60.5	223	2	US-08-615-271-9
19	701	60.5	223	3	US-09-074-660-9
20	701	60.5	223	3	US-09-074-659-9
21	701	60.5	223	4	US-09-106-468-9
22	701	60.5	223	4	US-09-106-468-9
23	701	60.5	223	3	US-09-120-592-2
24	701	60.5	223	3	US-09-120-592-2
25	701	60.5	223	4	US-08-944-483-35
26	691	59.7	224	4	US-08-944-483-35
27					

28	661.5	57.1	281	1	US-08-467-155A-7	Sequence 7, Appli
29	661.5	57.1	281	2	US-08-628-198-7	Sequence 7, Appli
30	661.5	57.1	281	3	US-09-201-038-7	Sequence 7, Appli
31	661.5	57.1	281	4	PCT-US96-07343-7	Sequence 7, Appli
32	619.5	53.5	299	1	US-08-978-4048-44	Sequence 8, Appli
33	619.5	53.5	299	2	US-08-944-483-34	Sequence 8, Appli
34	619.5	53.5	299	4	US-09-236-219-13	Sequence 8, Appli
35	619.5	53.5	299	5	PCT-US96-07343-8	Sequence 8, Appli
36	549.5	47.5	286	1	US-08-467-155A-9	Sequence 9, Appli
37	549.5	47.5	286	2	US-08-628-198-9	Sequence 9, Appli
38	549.5	47.5	286	4	US-09-201-038-9	Sequence 9, Appli
39	549.5	47.5	286	5	PCT-US96-07343-9	Sequence 9, Appli
40	486	42.0	156	4	US-09-261-416-6	Sequence 9, Appli
41	466	40.2	151	4	US-09-518-046-21	Sequence 21, Appli
42	451	38.9	282	3	US-09-025-059-1	Sequence 21, Appli
43	451	38.9	282	4	US-08-944-483-24	Sequence 24, Appli
44	449.5	38.9	282	4	US-08-386-642-14	Sequence 14, Appli
45	448	38.7	244	1	US-08-361-195-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-467-155A-10
; Sequence 10, Application US/08467155A
; Patent No. 5736377
; GENERAL INFORMATION:
; APPLICANT: Band, Vmla
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; TITLES OF SEQUENCES: 11
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30162
; REFERENCE/DOCKET NUMBER: 00398/100001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-155A-10

Query Match 71.1%; Score 823.5; DB 1; Length 271;
Best Local Similarity 63.4%; Pctid No. 2.2e-86;
Matches 168; Conservative 16; Mismatches 36; Indels 43; Gaps 7;
QY 1 IVGGY-----ECTSHSQVHSLNSCGSLNCGVSLNENWVSAACHQSRVEVLGE 52
DB 10 IVGGYSTRYPTECKKAYSQPHVSLNSCGVSLNENWVSAACHQSRVEVLGE 69

QY	1	IUGY-----ECTHSAOHVSLNSGVHXCGLSLINXWVVAACHYKSLVLRVLGE 53
DB	10	IUGVGSSTRPIIECKAYSQPHVSLNSGVHXCGLSLINWVVAACHYKSVRVLGE 69
QY	53	HH-----IRWNGTEQRISSSSVVRHPNTSSYNTXNDIMLKJUYDPTL----- 96
DB	70	WASGAEVETLVVGGSGFPISSRPVHPNYSSVNTINDIMLKLSKPAESTSTRPII 129

QY	1	I V G Y	-----	B C T H S Q A Q I S N S Y C H C C G S I N X K V V S A A C Y K S V L R V L G E	52
DB	10	I V G S Y S T R P I I E C A Y S Q I S Q I S N S Y T H C C G S I V N E N V V S A A C Y Q S R V E L R G E	69		
QY	53	H H	-----	I R N G E C T O X I S S S V S Y R P N Y S S Y N I K N D I M L I K L T P A T L	96
DB	70	H N S T R P I I I Q T G S E Q F I S S R V I R H P N Y S S Y N I D N I M L I K S P A T L S T R P I I	129		
QY	97	N O T H A V A L T E C A A D T M C Y S G W G N T M S S Y X D G D K Q X L K P I L S H A	148		

```

Db 130 NTYVQVVALPTSCAPACTWCTSGMNTNSTADKNKLOCLNIPILSYSSSTRYPIDCN 189
QY 149 NSYQPMITQSMFCAGYLEGKDCQSGSGPVVCGVGVWSGVT-----GCA 200
Db 190 NSY-PQMTNMFCAAGYLEGKDCQSGSGPVVCGELQ--GWSMGYSSTRYPIDCA 246
QY 201 ERDXPGVAKVAVLGSWVRDTMAY 225
Db 247 EPGNPGVAKVCFNDMLTSTMATY 271

RESULT 4
PCT-US96-07343-10
; Sequence 10, Application PC/TUS9607343
; GENERAL INFORMATION:
; APPLICANT: New England Medical Center Hospitals, Inc.
; TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
; MOLECULES OF SEQUENCES 11
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ben & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07343
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/467,155
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/100001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-07343-10

Query Match 71.1%; Score 823.5; DB 5; Length 271;
Best Local Similarity 63.4%; Pred. No. 2, 2e-86;
Matches 168; Conservative 16; Mismatches 38; Indels 43; Gaps 7;

QY 1 IVGGY-----ECTXHSQAQVSLNSGYHXCGLSLNKKXVWSAAHCYKSLVRLG 52
Db 10 IVGGYSTRYPIECAYSQPIQVSLNSGYHFCGGLSLNKKXVWSAAHCYKSLVRLG 69
QY 53 HH-----IRVNGETQKISSSVYRHPNYSYNDMLIKLTXPATI----- 96
Db 70 HNSSTRPIIIQVTEGSSQPISSSRVIRHPNYSYNDMLIKLSPATLSSTRPII 129
QY 97 NOYHVAALPTGACADATCTVSGKNTWSSVVDGDKLQXLPILSHA-----DCA 148
Db 130 NTYVQVVALPTSCAPACTWCTSGMNTNSTADKNKLOCLNIPILSYSSSTRYPIDCN 189
QY 149 NSYQPMITQSMFCAGYLEGKDCQSGSGPVVCGVGVWSGVT-----GCA 200

```

```

Db 190 NSY-PQMTNMFCAAGYLEGKDCQSGSGPVVCGELQ--GWSMGYSSTRYPIDCA 246
QY 201 ERDXPGVAKVAVLGSWVRDTMAY 225
Db 247 EPGNPGVAKVCFNDMLTSTMATY 271

RESULT 5
US-08-978-404B-44
; Sequence 44, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: FAST CELL PROTEASE THAT CLEAVES
; MOLECULES OF SEQUENCES 44
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5968782
US-08-978-404B-44

Query Match 65.2%; Score 755; DB 2; Length 246;
Best Local Similarity 63.8%; Pred. No. 1.4e-78;
Matches 143; Conservative 25; Mismatches 52; Indels 4; Gaps 3;

QY 1 IVGGYECTXHSQAQVSLNSGYHXCGLSLNKKXVWSAAHCYKSLVRLGHEHIFVNSG 60
Db 24 IVGGYTCTPESHVPQVSLNSGYHFCGGLSLNKKXVWSAAHCYKSLVRLGHEHIFVNSG 83
QY 61 TEQKISSSVYRHPNYSYNDMLIKLTXPATILOYHVALPTGACADATCTVSG 120
Db 84 DEQFINAKKIIKHPNYSYNDMLIKLSPVIRNARVAVLPSACAPAGTCLISG 143
QY 121 WNTWSS-VYDGDKLOXLPILSHADCANYSYSGMTQSMFCAGYLEGKDCQSGSG 179
Db 144 NGNTLSGVNPNLLQVDAPVLSQDCEAY-PGEITSSMTCVGLFEGKDCQSGSG 202
QY 180 PVYCVGLQGVVSGVWGCAERDYPGVAKVAVLGSWVRDTMA 223
Db 203 PVYCVGQLQ--GVSMGYGCAFPDNPSTVTKVCFNVGMQIDTIA 244

```

RESULT 6
 US-08-944-483-34
 ; Sequence 34, Application US/08944483
 ; Patent No. 6232456
 ; GENERAL INFORMATION:
 ; APPLICANT: COHEN, NAURICE
 ; INVENTOR: COHEN, NAURICE
 ; APPLICANT: FIEDMAN, PAULA N.
 ; INVENTOR: FIEDMAN, PAULA N.
 ; APPLICANT: GRADDOO, EDWARD N.
 ; INVENTOR: GRADDOO, EDWARD N.
 ; APPLICANT: KASS, MICHAEL R.
 ; INVENTOR: KASS, MICHAEL R.
 ; APPLICANT: RUSSELL, JOHN C.
 ; INVENTOR: RUSSELL, JOHN C.
 ; APPLICANT: STEWART, KENT D.
 ; INVENTOR: STEWART, KENT D.
 ; APPLICANT: STROUPE, STEVEN D.
 ; INVENTOR: STROUPE, STEVEN D.
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
 ; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
 ; TITLE OF INVENTION: OF THE PROSTATE
 ; NUMBER OF SEQUENCES: 76
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Abbott Laboratories
 ; STREET: 100 Abbott Park Road
 ; CITY: Abbott Park
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60064-3500
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/944,483
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Becker, Cheryl L.
 ; REGISTRATION NUMBER: 35,441
 ; REFERENCE/DOCKET NUMBER: 6183 US.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 847/935-1729
 ; TELEFAX: 847/935-4623
 ; FAX:
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 224 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: No. 6232456
 ; US-08-944-483-34
 ;
 Query Match 61.3%; Score 710; DB 4; Length 224;
 Best Local Similarity 60.3%; Pred. No. 1.7e-73;
 Matches 135; Conservative 31; Mismatches 54; Indels 4; Gaps 3;
 QY 1 IVGGVECTXHSQAHQVSLNSGYHXCXGSLINXWVYSAACHYKSVLRVRLGHEHIVNEG 60
 DB 1 IVGGVCEINSVPYQVSLNSGYHFCXGSLISEQWVYVSAHGYKSVLRVRLGHEHIVNEG 60
 QY 61 TEQXISSSSVXKHPNYSYXNDIMLIKLTYPATLNQYVHAVALPTECAADATMCTVSG 120
 DB 61 NEQFTINAKIIRHPKYNSTLNDLILIKLSSPAVINSRVSALSTPPAAGTESLISG 120
 QY 121 WGMTMSSVXD-GDKLQXLPILSHADCANSGYGMITQSMPCAGYLEGKSCQDSGG 179
 DB 121 WGMTLSSGADYDPELOCLDAPVLSQAECASY-PGKITNNPFCVGLGKSCQDSGG 179
 QY 180 PWCNGLVQGVWGWGCAERDXPQYAKYXVLSGWWRDTMA 223
 DB 180 PWSNGELQ--GIVSWGYGCAQKRRPGVTKYVYVDMIKDTIA 221

RESULT 7
 US-08-956-267A-2
 ; Sequence 2, Application US/08956267A
 ; Patent No. 5945328
 ; GENERAL INFORMATION:
 ; APPLICANT: WOLDIKE, Helle Fabricius
 ; INVENTOR: WOLDIKE, Helle Fabricius
 ; APPLICANT: KJELDSEN, Thomas Borglum
 ; INVENTOR: KJELDSEN, Thomas Borglum
 ; TITLE OF INVENTION: A Process For Producing Trypsin
 ; TITLE OF INVENTION: (Trypsinogen)
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 59453280 No. 5945328diak of No. 5945328th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/956,267A
 ; FILING DATE: 22-OCT-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ROSEN, GERT
 ; REGISTRATION NUMBER: 36,993
 ; REFERENCE/DOCKET NUMBER: 4500-204-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 247 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-956-267A-2
 ;
 Query Match 61.1%; Score 707; DB 2; Length 247;
 Best Local Similarity 60.3%; Pred. No. 4.4e-73;
 Matches 135; Conservative 27; Mismatches 58; Indels 4; Gaps 3;
 QY 1 IVGGVECTXHSQAHQVSLNSGYHXCXGSLINXWVYSAACHYKSVLRVRLGHEHIVNEG 60
 DB 25 IVGGVTCANSPYQVSLNSGYHFCXGSLINQWVYSAACHYKSVLRVRLGHEHIVNEG 84
 QY 61 TEQXISSSSVXKHPNYSYXNDIMLIKLTYPATLNQYVHAVALPTECAADATMCTVSG 120
 DB 85 NEQFTINAKIIRHPKYNSTLNDLILIKLSSPAVINSRVSALSTPPAAGTESLISG 144
 QY 121 WGMTMSSVXD-GDKLQXLPILSHADCANSGYGMITQSMPCAGYLEGKSCQDSGG 179
 DB 145 WGMTKSSGSSVPYSLQCLKAPVLSDDSKSSY-PQKITNNMICVGLGKSCQDSGG 203
 QY 180 PWCNGLVQGVWGWGCAERDXPQYAKYXVLSGWWRDTMA 223
 DB 204 PWCNGQLQ--GIVSWGYGCAQKRRPGVTKYVYVDMIKDTIA 245
 ;
 RESULT 8
 US-08-766-982-13
 ; Sequence 13, Application US/08766982
 ; Patent No. 5948891
 ; GENERAL INFORMATION:
 ; APPLICANT: NABY, Robert C.
 ; INVENTOR: NABY, Robert C.
 ; TITLE OF INVENTION: Analogs of Macrophage Stimulating
 ; TITLE OF INVENTION: Protein
 ; NUMBER OF SEQUENCES: 13

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Inc.
STREET: 1840 Behavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,982
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-441
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-766-982-13

Query Match 60.6%; Score 702; DB 2; Length 224;
Best Local Similarity 60.3%; Pred. No. 1.4e-72;
Matches 135; Conservative 26; Mismatches 59; Indels 4; Gaps 3;

QY 1 IVGGYECTHSQAHQVSLNSGYHXCGLSLNXXVYSAACHYKSVLRLGHEHHRVNEG 60
Db 1 IVGGYECNEENSPVQVSLNSGYHFCGGLSLNEQWVSAAGHYKSRIQVRLGHEHNEVLEG 60

QY 61 TEQVSSSVYKHRRHNTSYNNXNDIMLIKLTXPATLNYQVHAVALPTECAADATMCTVSG 120
Db 61 NEQFINAAKIIRHPOYDRKTLNNDIMLIKLSRAVINARVSTISLTPPATGTCLISG 120

QY 121 WGMTASSYVD-GDKLQXKLPLSHADCANSGYGMITOSMFCAGYLEGGKSCQGDSSG 179
Db 121 WGMTASSGADYPDELQCLDAPVLSQAKCEASY-PGKITSNMFVGLGGKSCQGDSSG 179

QY 180 PIVCGVGLQGVWSGVCGERDYPGVYKXVYLSGWRDTMA 223
Db 180 PIVCGGQLQ-GVWSGDCGCAQKPGVYTKVYVYKVIQTIA 221

RESULT 9
US-08-944-483-36
Sequence 36, Application US/08944483
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: RUSSELL, MICHAEL R.
APPLICANT: STEWART, KEVIN C.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELETYPE:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-36

Query Match 60.6%; Score 702; DB 4; Length 224;
Best Local Similarity 60.3%; Pred. No. 1.4e-72;
Matches 135; Conservative 26; Mismatches 59; Indels 4; Gaps 3;

QY 1 IVGGYECTHSQAHQVSLNSGYHXCGLSLNXXVYSAACHYKSVLRLGHEHHRVNEG 60
Db 1 IVGGYECNEENSPVQVSLNSGYHFCGGLSLNEQWVSAAGHYKSRIQVRLGHEHNEVLEG 60

QY 61 TEQVSSSVYKHRRHNTSYNNXNDIMLIKLTXPATLNYQVHAVALPTECAADATMCTVSG 120
Db 61 NEQFINAAKIIRHPOYDRKTLNNDIMLIKLSRAVINARVSTISLTPPATGTCLISG 120

QY 121 WGMTASSYVD-GDKLQXKLPLSHADCANSGYGMITOSMFCAGYLEGGKSCQGDSSG 179
Db 121 WGMTASSGADYPDELQCLDAPVLSQAKCEASY-PGKITSNMFVGLGGKSCQGDSSG 179

QY 180 PIVCGVGLQGVWSGVCGERDYPGVYKXVYLSGWRDTMA 223
Db 180 PIVCGGQLQ-GVWSGDCGCAQKPGVYTKVYVYKVIQTIA 221

RESULT 10
US-08-946-219-13
Sequence 13, Application US/09296219
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: Wahl, Robert C.
TITLE OF INVENTION: Analogs of Macrophage Stimulating
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Inc.
STREET: 1840 Behavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/296,219

```

```

; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wincer, Ronald B.
; REFERENCE/DOC# NUMBER: A-441
; INFORMATION/DOC# ID NO: 13
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-236-219-13

Query Match 60.6%; Score 702; DB 4; Length 224;
Best Local Similarity 60.3%; Pred. No. 1.4e-72;
Matches 135; Conservative 26; Mismatches 59; Indels 4; Gaps 3;

Qy 1 IVGGVETXHSQAHQVSLNSGYHXCGLSLINXWVYSAHCYKSVLRVLGEHHRVNEG 60
Db 1 IVGGYCNCENSVPYQVSLNSGYHFCGSLINEQWVYSAGHCYKSVLRVLGEHHRVNEG 60
Qy 61 TEQXISSSSVXKHPNYSYNIANDIMLIKTXPATLNQYHVALPTECAADATMCTVSG 120
Db 61 NEQFINAAKIIRHPQYDRKTLNNDIMLIKLSRAVINARVSTISLPTAPPATGKCLISG 120
Qy 121 WGMTSSVXD-GDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGG 179
Db 121 WGMTASSGADYPDELQCLDAPVLSQACEASY-PGKITSNMFVGLFEGGKSCQDSGG 179
Qy 180 PVVCGVLQGVVSWNGYGCABRDXPGYAKVXVLSGWRDTMA 223
Db 180 PVVCGQLQ--GVVSWGDCQAQKPGVYTKVYVYVVKMIKNTIA 221

RESULT 11
US-09-027-337-5
; Sequence 5, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotsuhi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; Tissue of Invention: Breast and Ovarian Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/027,337B
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 5
; LENGTH: 225
; TYPE: PEPT
; ORGANISM: Unknown
; OTHER INFORMATION: Serine protease catalytic domain of trypsin (Try)
; OTHER INFORMATION: homologous to similar domain in TADG-15
US-09-027-337-5

Query Match 60.6%; Score 702; DB 2; Length 225;
Best Local Similarity 60.3%; Pred. No. 1.4e-72;
Matches 135; Conservative 26; Mismatches 59; Indels 4; Gaps 3;

Qy 1 IVGGVETXHSQAHQVSLNSGYHXCGLSLINXWVYSAHCYKSVLRVLGEHHRVNEG 60
Db 2 IVGGYCNCENSVPYQVSLNSGYHFCGSLINEQWVYSAGHCYKSVLRVLGEHHRVNEG 61
Qy 61 TEQXISSSSVXKHPNYSYNIANDIMLIKTXPATLNQYHVALPTECAADATMCTVSG 120
Db 62 NEQFINAAKIIRHPQYDRKTLNNDIMLIKLSRAVINARVSTISLPTAPPATGKCLISG 121
Qy 121 WGMTSSVXD-GDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGG 179
Db 122 WGMTASSGADYPDELQCLDAPVLSQACEASY-PGKITSNMFVGLFEGGKSCQDSGG 180
Qy 180 PVVCGVLQGVVSWNGYGCABRDXPGYAKVXVLSGWRDTMA 223
Db 181 PVVCGQLQ--GVVSWGDCQAQKPGVYTKVYVYVVKMIKNTIA 222

RESULT 13
US-08-278-091-9
; Sequence 9, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Fele
; APPLICANT: COHEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
; Tissue of Invention: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sam McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/278,091
/ FILING DATE: 21-JUL-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stewart, Michael I.
/ REGISTRATION NUMBER: 24,973
/ REFERENCE/DOCKET NUMBER: 1038-371
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 595-1155
/ TELEFAX: (416) 595-1163
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 223 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ US-08-278-091-9
/
/ Query Match 60.5%; Score 701; DB 1; Length 223;
/ Best Local Similarity 60.3%; Pred. No. 1.8e-72;
/ Matches 135; Conservative 29; Mismatches 56; Indels 4; Gaps 3;
/
/ Qy 1 IVGGYECTHSAQHVSLSNGYHKCGSLINXWVSAHCYKSVLRVLRGEHHIRVNEG 60
/ Db 1 IVGGYTCGANTVPQVLSNGYHFCGSLINSQWVSAHCYKSGIOVRLGEDINIVVEG 60
/
/ Qy 61 TEQXISSSVXKHPNYSYXNDIMLIKTXPATLNQYHVAVALPTECAADATMCTVSG 120
/ Db 61 NEQFISAKSI VHPSYNSNTLNNDIMLIKSAASLSNRVASISLPTSCASAGTQCLISG 120
/
/ Qy 121 WGNTHSS-VVGDGKLQXLPILSHADCANSGYGMITQSMFCAGYLEGKXDCQDGGG 179
/ Db 121 WGNTHSS-VVGDGKLQXLPILSHADCANSGYGMITQSMFCAGYLEGKXDCQDGGG 179
/
/ Qy 180 PWCNGLVQGVWSNGYCAERDPGVYAKVYLSGWRDTMA 223
/ Db 180 PWCNGLVQGVWSNGYCAERDPGVYAKVYLSGWRDTMA 223
/
/ RESULT 15
/ US-08-472-173-9
/ Sequence 9, Application US/08472173
/ Patent No. 5665353
/ GENERAL INFORMATION:
/ APPLICANT: LOOSMORE, Sheena M.
/ APPLICANT: YANG, Yan-Ping
/ APPLICANT: CHONG, Pele
/ APPLICANT: CHONG, Pele
/ APPLICANT: KLEIN, Michael H.
/ APPLICANT: KLEIN, Michael H.
/ TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
/ TITLE OF INVENTION: Reduced Protease Activity
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: Suite 701, 330 University Avenue
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/472,173
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/296,149
/ FILING DATE: 26-AUG-1994
/
/ Qy 1 IVGGYECTHSAQHVSLSNGYHKCGSLINXWVSAHCYKSVLRVLRGEHHIRVNEG 60
/ Db 1 IVGGYTCGANTVPQVLSNGYHFCGSLINSQWVSAHCYKSGIOVRLGEDINIVVEG 60
/
/ Qy 61 TEQXISSSVXKHPNYSYXNDIMLIKTXPATLNQYHVAVALPTECAADATMCTVSG 120
/ Db 61 NEQFISAKSI VHPSYNSNTLNNDIMLIKSAASLSNRVASISLPTSCASAGTQCLISG 120
/
/ Qy 121 WGNTHSS-VVGDGKLQXLPILSHADCANSGYGMITQSMFCAGYLEGKXDCQDGGG 179
/ Db 121 WGNTHSS-VVGDGKLQXLPILSHADCANSGYGMITQSMFCAGYLEGKXDCQDGGG 179
/
/ Qy 180 PWCNGLVQGVWSNGYCAERDPGVYAKVYLSGWRDTMA 223
/ Db 180 PWCNGLVQGVWSNGYCAERDPGVYAKVYLSGWRDTMA 223
/
/ RESULT 14
/ US-08-483-859-9
/ Sequence 9, Application US/08483859
/ Patent No. 5656436
/ GENERAL INFORMATION:
/ APPLICANT: LOOSMORE, Sheena M.
/ APPLICANT: YANG, Yan-Ping
/ APPLICANT: CHONG, Pele
/ APPLICANT: CHONG, Pele
/ APPLICANT: KLEIN, Michael H.
/ APPLICANT: KLEIN, Michael H.
/ TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
/ TITLE OF INVENTION: Reduced Protease Activity
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sim & McBurney
/ STREET: Suite 701, 330 University Avenue
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5G 1R7
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/483,859
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:WG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-173-9

Query Match          60.5%; Score 701; DB 1; Length 223;
Best Local Similarity 60.3%; Pred. NO. 1.8e-72;
Matches 135; Conservative 29; Mismatches 56; Indels 4; Gaps 3;

Qy 1 IVGGYECTXSHQHVSLNSGYHKCGSLINXXWVSAAHCKYKSLVRLGEHHIRVNEG 60
Db 1 IVGGYTCGANTVPYQVSLNSGYHFCGSLNSQWVSAAHCKYKSGIQVRLGEDINVVEG 60

Qy 61 TEQXISSSVARHPNYSSYINXNDIMLIKLTXPATLNYVYHVALPTECAADATMCTVSG 120
Db 61 NEQFISAKSVHPSYNSNTLNNDIMLIKLGAASLSNRVASISLPTSCAGAGTQCLISG 120

Qy 121 WGNMTSS-VXDGNKLQXILXLPILSHADCANSGYRPMITQSMFCAGYLEGGKSCGGSGG 179
Db 121 WGNKXSGTGYPDVLKCLKAPILSDSCSKAY-PQGITSMFCAGYLEGGKSCGGSGG 179

Qy 180 PWCNGVVLQGVVWVWSGYGCAERDYPGYAKVXVLSCWVRDTMA 223
Db 180 PWCGRKLQ--GIVSWGCGCAQKRPGYITKCNVSVWIKQTIA 221

Search completed: February 12, 2003, 10:30:12
Job time : 49.3562 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:17:50 ; Search time 56.4179 Seconds
(without alignments)
383.393 Million cell updates/sec

Title: US-10-036-371-8

Perfect score: 1158

Sequence: 1 IVGGECTXHSQAHQVLSNS.....GVYAKVXVLSGVRDTMAXY 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 43 summaries

Database :

PIR 73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086.5	93.8	240	2 S31947	trypsin (EC 3.4.21
2	1085.5	93.7	241	2 S39048	trypsin (EC 3.4.21
3	905.5	78.2	242	2 S31776	trypsin (EC 3.4.21
4	901.5	77.8	242	2 S31775	trypsin (EC 3.4.21
5	897.5	77.5	242	2 S49489	trypsin (EC 3.4.21
6	893.5	77.2	231	2 S31778	trypsin (EC 3.4.21
7	758.5	65.5	229	1 TRDPS	trypsin (EC 3.4.21
8	756	65.3	238	2 S31779	trypsin (EC 3.4.21
9	755	65.2	246	1 TRR11	trypsin (EC 3.4.21
10	729	63.0	246	2 B25528	trypsin (EC 3.4.21
11	727	62.8	247	1 TRTDG	trypsin (EC 3.4.21
12	726	62.7	246	1 TRRT2	trypsin (EC 3.4.21
13	721	62.3	247	2 S31813	trypsin (EC 3.4.21
14	718	62.1	247	2 A35871	trypsin (EC 3.4.21
15	716	62.0	243	2 TRB572	trypsin (EC 3.4.21
16	710	61.9	243	1 TRB572	trypsin (EC 3.4.21
17	707	61.1	231	1 TRDQC	trypsin (EC 3.4.21
18	706	61.0	246	1 TRDQC	trypsin (EC 3.4.21
19	702	60.6	247	1 A25952	trypsin (EC 3.4.21
20	701	60.5	229	1 TRB07R	trypsin (EC 3.4.21
21	696	60.1	248	2 S55067	trypsin (EC 3.4.21
22	691	59.7	304	2 S33496	trypsin (EC 3.4.21
23	681.5	58.9	259	2 S38363	trypsin (EC 3.4.21
24	674	58.2	248	2 S55066	trypsin (EC 3.4.21
25	665	57.4	247	2 S12764	trypsin (EC 3.4.21
26	650.5	56.2	247	2 S05494	trypsin (EC 3.4.21
27	617.5	53.3	246	2 QJ1472	trypsin (EC 3.4.21
28	616.5	53.2	246	2 QJ1471	trypsin (EC 3.4.21
29	566	48.9	250	2 TQ1779	trypsin (EC 3.4.21

trypsin (EC 3.4.21
trypsin (EC 3.4.21
trypsin (EC 3.4.21
trypsin (EC 3.4.21
trypsin (EC 3.4.21
trypsin (EC 3.4.21
chymotrypsin (EC 3
chymotrypsin (EC 3
7S nerve growth fa
tissue kallikrein
tissue kallikrein
tissue kallikrein
chymotrypsin B - A
trypsin (EC 3.4.21
serine proteinase
trypsin-like prote
tissue kallikrein

ALIGNMENTS

RESULT 1

S39047

trypsin (EC 3.4.21.4) I - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000

C:Accession: S39047

Eur. J. Biochem. 217, 1091-1097, 1993

A:Title: Isolation and characterization of cDNAs from Atlantic cod encoding two different

A:Reference number: S39047; MUID:94039130; PMID:8223632

A:Postulate type: mRNA

A:Cross references: EMBL:X7686; NID:9450517; PIDN:CM54214.1; PID:g1334752

C:Superfamily: trypsin; trypsin homolog

C:Keywords: hydrolase; trypsin; serine proteinase

F:19-233/Domain: trypsin homology <TRY>

F:56,102,194/Active site: His, Asp, Ser #status predicted

Query Match

Best Local Similarity 93.8%; Score 1086.5; DB 2; Length 240;

Matches 206; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

Qy 1 IVGGECTXHSQAHQVLSNSGYHXCGSLINXVWSAAHCYKSVLRVLGEHHRVNEG 60

Db 19 IVGGECTXHSQAHQVLSNSGYHXCGSLINXVWSAAHCYKSVLRVLGEHHRVNEG 78

Qy 61 TEQYISSSVYHFNYSINXNDMLIKITPATLNYVHAVALPTECADATVCTVSG 120

Db 79 TEQYISSSVYHFNYSINXNDMLIKITPATLNYVHAVALPTECADATVCTVSG 138

Qy 121 MONTSSVYDGLKQLXLPITLSHADCANSGPOMTOSMFCAGYLEGKSGCGSGCP 180

Db 139 MONTSSVYDGLKQLXLPITLSHADCANSY-POMTOSMFCAGYLEGKSGCGSGCP 197

Qy 181 VVCNGLQGVGVSGVGGCAERDGVYAKVXVLSGVRDTMAXY 225

Db 198 VVCNGLQ--GVVSGVGGCAERDGVYAKVXVLSGVRDTMAXY 240

RESULT 2

S39048

trypsin (EC 3.4.21.4) X - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 20-May-1994 #sequence_revision 03-Aug-1995 #text_change 20-Jun-2000

C:Accession: S39048

Eur. J. Biochem. 217, 1091-1097, 1993

A:Title: Isolation and characterization of cDNAs from Atlantic cod encoding two different

A:Reference number: S39047; MUID:94039130; PMID:8223632

A:Postulate type: preliminary

RESULT 4
S31775
trypsin (EC 3.4.21.4) I precursor - Atlantic salmon
C:Species: Salmo salar (Atlantic salmon)
C>Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 15-Oct-1999
C:Accession: S66650; S66661; S31775; S31777
C:Keywords: hydrolyase; serine proteinase
P:20,234/Domain: trypsin homology <TRY>
P:59,103,195/Active site: His, Asp, Ser #status predicted

Query Match 93.7%; Score 1085.5; DB 2; Length 241;
Best Local Similarity 91.6%; Pred. No. 6.2e-99;
Matches 206; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

QY 1 IVGGVETCTHSAHQVSLNSGYHFCGSLINXVWVSAACHYKSVLRVLGHEHHRVNEG 60
DB 20 IVGGVETCTHSAHQVSLNSGYHFCGSLVWVWVSAACHYKSVLRVLGHEHHRVNEG 79
QY 61 TEQXISSSSVKRRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 120
DB 80 TEQPTSSSRVIRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 139
QY 121 WGMTMSSVXDKDLQKLXLPILSHADCANSGYGMITQSMFCAGYLEGKDCSQGDSGGP 180
DB 140 WGMTMSSVDDGDKLQCLNLPILSHADCANSGY-PGMITQSMFCAGYLEGKDCSQGDSGGP 199
QY 181 VVCGVLTQGVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 225
DB 199 VVCGVLTQ--GVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 241

Query Match 78.2%; Score 905.5; DB 2; Length 242;
Best Local Similarity 75.6%; Pred. No. 2.8e-81;
Matches 170; Conservative 15; Mismatches 37; Indels 3; Gaps 2;

QY 1 IVGGVETCTHSAHQVSLNSGYHFCGSLINXVWVSAACHYKSVLRVLGHEHHRVNEG 60
DB 21 IVGGVETCTHSAHQVSLNSGYHFCGSLVWVWVSAACHYKSVLRVLGHEHHRVNEG 80
QY 61 TEQXISSSSVKRRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 120
DB 81 TEQPTSSSRVIRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 140
QY 121 WGMTMSSVXDKDLQKLXLPILSHADCANSGYGMITQSMFCAGYLEGKDCSQGDSGGP 180
DB 141 WGMTMSSVDDGDKLQCLNLPILSHADCANSGY-PGMITQSMFCAGYLEGKDCSQGDSGGP 199
QY 181 VVCGVLTQGVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 225
DB 200 VVCGVLTQ--GVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 242

RESULT 4
S31775
trypsin (EC 3.4.21.4) I precursor - Atlantic salmon
C:Species: Salmo salar (Atlantic salmon)
C>Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 15-Oct-1999
C:Accession: S66650; S66661; S31775; S31777
C:Keywords: hydrolyase; serine proteinase
P:20,234/Domain: trypsin homology <TRY>
P:59,103,195/Active site: His, Asp, Ser #status predicted

Query Match 93.7%; Score 1085.5; DB 2; Length 241;
Best Local Similarity 91.6%; Pred. No. 6.2e-99;
Matches 206; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

QY 1 IVGGVETCTHSAHQVSLNSGYHFCGSLINXVWVSAACHYKSVLRVLGHEHHRVNEG 60
DB 20 IVGGVETCTHSAHQVSLNSGYHFCGSLVWVWVSAACHYKSVLRVLGHEHHRVNEG 79
QY 61 TEQXISSSSVKRRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 120
DB 80 TEQPTSSSRVIRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 139
QY 121 WGMTMSSVXDKDLQKLXLPILSHADCANSGYGMITQSMFCAGYLEGKDCSQGDSGGP 180
DB 140 WGMTMSSVDDGDKLQCLNLPILSHADCANSGY-PGMITQSMFCAGYLEGKDCSQGDSGGP 199
QY 181 VVCGVLTQGVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 225
DB 199 VVCGVLTQ--GVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 241

Query Match 77.5%; Score 897.5; DB 2; Length 242;
Best Local Similarity 75.1%; Pred. No. 6.8e-81;
Matches 169; Conservative 15; Mismatches 36; Indels 3; Gaps 2;

QY 1 IVGGVETCTHSAHQVSLNSGYHFCGSLINXVWVSAACHYKSVLRVLGHEHHRVNEG 60
DB 21 IVGGVETCTHSAHQVSLNSGYHFCGSLVWVWVSAACHYKSVLRVLGHEHHRVNEG 80
QY 61 TEQXISSSSVKRRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 120
DB 81 TEQPTSSSRVIRHPNYSYNDIMLKLKTPATLNQYHVAVALPTCEADATMCTVSG 140
QY 121 WGMTMSSVXDKDLQKLXLPILSHADCANSGYGMITQSMFCAGYLEGKDCSQGDSGGP 180
DB 141 WGMTMSSVDDGDKLQCLNLPILSHADCANSGY-PGMITQSMFCAGYLEGKDCSQGDSGGP 199
QY 181 VVCGVLTQGVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 225
DB 200 VVCGVLTQ--GVVSGVYGCARDXPGVYAKVYKVLGWRVDTMXY 242

RESULT 5
S49489
trypsin (EC 3.4.21.4) precursor - Paranotothenia magellanica
C:Species: Paranotothenia magellanica
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C:Accession: S49489
R:Genicot, S.; Rentier-Delrue, F.; Edwards, D.; van Beumen, J.; Dodson, G.; Gerday, C.
submitted to the EMBL Data Library, October 1994
A:Description: Trypsin and trypsinogen from an antarctic fish: molecular basis of cold ac-
tivation
A:Reference number: S49489
A:Accession: S49489
A:Molecule type: mRNA
A:Residues: 1-242 <GEN>
A:Cross-references: EMBL:X82223; NID:9559507; PID:CA547701.1; PID:9559508
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolyase; serine proteinase
P:1-13/Domain: signal sequence #status predicted <SIG>
P:124-242/Product: trypsin #status predicted <MAT>
P:121-235/Domain: trypsin homology <TRY>

```

Best Local Similarity 75.1%; Pred. No. 1.7e-80;
Matches 169; Conservative 16; Mismatches 37; Indels 3; Gaps 2;

QY 1 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 60
Db 21 IVGGKCSYSPHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 80
QY 61 TEQXISSSVVRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 120
Db 81 KEQFISSSVRIRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 140
QY 121 WNTNSVSVGDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGGP 180
Db 141 WGSTOSSADGNKLOCLNLPILSRDCDSNY-PGMITDAMFCGLQSGKUSCQDSGGP 199
QY 181 VVCGNVLQGVGVNSGYGCAERDPCGYAKYVLSGWVDRTMAY 225
Db 200 VVCGNELQ--GVWSNGYGCARDHPGYAKYVLSGWVDRTMAY 242

RESULT 6
S31778
C:Species: Salmo gairdneri - Atlantic salmon (fragment)
C:Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S66657; S31779
R:Male, R.; Lorenz, J.B.; Smalas, A.O.; Torriassen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of trypsin
A:Reference number: S66657; MUID:96035908; PMID:7556223
A:Accession: S66658
A:Molecule type: mRNA
A:Residues: 1-231 <NAL>
A:Cross-references: EMBL:X70073; NID:964385; PID:CAA9678.1; PID:964386
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
P:1-7/Domain: activation peptide [fragment] #status predicted <SIG>
P:8-229/Product: trypsin #status predicted <MAT>
P:10-231/Domain: trypsin homology <TRY>
P:12-244/Domain: trypsin homology <TRY>
P:140-244/Domain: trypsin homology <TRY>
P:149,93,185/Active site: His, Asp, Ser #status predicted
P:74,9,93,185/Active site: His, Asp, Ser #status predicted

Query Watch 77.2%; Score 893.5; DB 2; Length 231;
Best Local Similarity 74.7%; Pred. No. 3.9e-80;
Matches 168; Conservative 16; Mismatches 36; Indels 3; Gaps 2;

QY 1 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 60
Db 10 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 80
QY 61 TEQXISSSVVRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 120
Db 70 SEQFISSSVRIRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 129
QY 121 WNTNSVSVGDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGGP 180
Db 130 WNTNSVSVGDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGGP 188
QY 181 VVCGNVLQGVGVNSGYGCAERDPCGYAKYVLSGWVDRTMAY 225
Db 189 VVCGNELQ--GVWSNGYGCARDHPGYAKYVLSGWVDRTMAY 231

RESULT 7
TRDFS
C:Species: Salmo gairdneri - spiny dogfish
C:Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 21-Jan-1997
C:Accession: A00950; B27719
R:Itani, K.; Britson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975

```

```

A:Title: Amino acid sequence of dogfish trypsin.
A:Reference number: A00950; MUID:75146445; PMID:1092332
A:Accession: A00950
A:Molecule type: Protein
A:Residues: 229
A:Note: 119-Pro was also found
R:Hermodson, M.A.; Tyler, R.W.; Reek, G.R.; Neurath, H.; Walsh, K.A.
FEBS Lett. 14, 222-224, 1971.
A:Title: Comparison of the amino terminal sequences of bovine, dogfish, and lungfish trypt
A:Reference number: B27719
A:Accession: B27719
A:Molecule type: protein
A:Residues: 1-21 <HER>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; sym
P:1-7/Domain: activation peptide #status experimental <APT>
P:8-229/Product: trypsin #status predicted <MAT>
P:10-244/Domain: trypsin homology <TRY>
P:12-244/Domain: trypsin homology <TRY>
P:140-244/Domain: trypsin homology <TRY>
P:149,93,185/Active site: His, Asp, Ser #status predicted
P:74,9,93,185/Active site: His, Asp, Ser #status predicted
P:59,61,64,69/Binding site: calcium (Glu, Asp, Ala, Glu) #status predicted

Query Watch 65.5%; Score 758.5; DB 1; Length 229;
Best Local Similarity 62.8%; Pred. No. 6.7e-67;
Matches 140; Conservative 23; Mismatches 57; Indels 3; Gaps 2;

QY 1 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 60
Db 8 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 67
QY 61 TEQXISSSVVRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 120
Db 68 DTDYSSVVRHPNYSYNIKNDIMLKLTXPATLNQYVHAVALPTECAADATMCTVSG 127
QY 121 WNTNSVSVGDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGGP 180
Db 128 WNTNSVSVGDKLQXLPILSHADCANSGPGMITQSMFCAGYLEGGKSCQDSGGP 186
QY 181 VVCGNVLQGVGVNSGYGCAERDPCGYAKYVLSGWVDRTMAY 223
Db 187 VVCGNELQ--GVWSNGYGCARDHPGYAKYVLSGWVDRTMAY 227

RESULT 8
S31779
C:Species: Salmo gairdneri - Atlantic salmon (fragment)
C:Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S66657; S31779
R:Male, R.; Lorenz, J.B.; Smalas, A.O.; Torriassen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of trypsin
A:Reference number: S66657; MUID:96035908; PMID:7556223
A:Accession: S66658
A:Molecule type: mRNA
A:Residues: 1-238 <NAL>
A:Cross-references: EMBL:X70074; NID:964387; PID:CAA9679.1; PID:964388
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
P:1-7/Domain: signal sequence [fragment] #status predicted <SIG>
P:8-15/Domain: activation peptide [fragment] #status predicted <APT>
P:16-238/Product: trypsin III #status predicted <MAT>
P:16-231/Domain: trypsin homology <TRY>
P:22-152,40-56,124-225,131-198,163-177,188-212/Disulfide bonds: #status predicted
P:55,99,192/Active site: His, Asp, Ser #status predicted

Query Watch 65.3%; Score 756; DB 2; Length 238;
Best Local Similarity 63.8%; Pred. No. 1.2e-66;
Matches 143; Conservative 26; Mismatches 51; Indels 4; Gaps 3;

QY 1 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 60
Db 16 IVGGVECTKXSOAHQVSLNSGYHCKGSLINXXWVSAACHYKSVLRVLGEHHIRVNEG 75

```

```

Qy 61 TEQXISSSVXRHPNYSYNIKNIMLIKLTXPATLNOYVHAVALPTFCADATMCTVSG 120
Db 76 TEQFIDSVKVMESYNSRLNDIMLIKUSPALSLSYVVALPSSCASGTRCLVSG 135
Qy 121 WGNMTSSVXD-GDKQLXKLPLTSLHADCAANSYGMITQSMFCAGYLEGGKDCQDSGG 179
Db 136 WGNLSGSSNYPFTLRCLDPLILSSSSCNAY-PQGITSSMFCAGYLEGGKDCQDSGG 194
Qy 180 PWCNGLQGVGWSNGYGCABNDXPGYAKVXVLSGWRDTMA 223
Db 195 PWCNGQLQ--GWSNGYGCABNDXPGYTKVNCYRWSISSTMS 236
RESULT 9
THRT1
trypsin (EC 3.4.21.4) I precursor - rat
N|Alternate names: trypsinogen I
C|Species: Rattus norvegicus (Norway rat)
C|Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999
R|Accession: B2657; GI: 59346; SWIFT, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
R|Catal. Chem. 257, 9724-9732, 1982
J|Biol. Chem. 259, 14255-14264, 1984
A|Title: Structure of two related rat pancreatic trypsin genes.
A|Reference number: A23657; MUID: 85054880; PMID: 6094547
A|Accession: B22657
A|Molecule type: DNA
A|Residues: 1-246 <CR>
A|Cross-references: GB:J00778; NID:9206507; PIDN:AAA98518.1; PID:g206508
A|Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17
R|MacDonald, R.J.; Stery, S.J.; Swift, G.H.
J|Biol. Chem. 257, 9724-9732, 1982
A|Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
A|Reference number: A00948; MUID: 82265624; PMID: 6896710
A|Accession: A00948
A|Molecule type: mRNA
A|Residues: 1-246 <MAC>
A|Cross-references: GB:J00778; NID:9206507; PIDN:AAA98518.1; PID:g206508
C|Genetics:
A|Introns: 14/1; 67/2; 152/1; 197/3
C|Superfamily: trypsin; trypsin homology
C|Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F|1-15/Domain: signal sequence #status predicted <SIG>
F|16-23/Domain: activation peptide #status predicted <APT>
F|24-246/Product: trypsin I #status predicted <EN2>
F|24-239/Domain: trypsin homology <TRY>
F|30-160,48-64,132-233,139-206,171-185/Diulfide bonds: #status predicted
F|63,107,200/Active site: His, Asp, Ser #status predicted
F|75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 65.2%; Score 755; DB 1; Length 246;
Best Local Similarity 63.8%; Pred. No. 1.6e-66;
Matches 143; Conservative 29; Mismatches 52; Indels 4; Gaps 3;

Qy 1 IVGGYECTHSQAHQVSLNSGVKXGSLINXWVSAACHYKSVLRVLRGHHIRVNEG 60
Db 24 IVGGTTPESHVPYQVSLNSGVHFCGSLINDQVWVSAACHYKSRIVQLRGEHNIVLEG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNIMLIKLTXPATLNOYVHAVALPTFCADATMCTVSG 120
Db 84 DEQFINAAKIIRPNYSMTLNDIMLIKUSPVKLNRVAVPALSPACAPAGTQCLISG 143
Qy 121 WGNMTSS--VADGDKLXKLPLTSLHADCAANSYGMITQSMFCAGYLEGGKDCQDSGG 179
Db 144 WGNLTSLNGVNPDLQCVDAFVPLSQADCEAA--PGEITSSMFCVGLSGKDCQDSGG 202
Qy 180 PWCNGLQGVGWSNGYGCABNDXPGYAKVXVLSGWRDTMA 223
Db 203 PWCNGQLQ--GIVSWGVCALPNDPFGVTKVNFVGMITQTTA 244
Query Match 65.2%; Score 755; DB 1; Length 246;
Best Local Similarity 63.8%; Pred. No. 1.6e-66;
Matches 143; Conservative 29; Mismatches 52; Indels 4; Gaps 3;

Qy 1 IVGGYECTHSQAHQVSLNSGVKXGSLINXWVSAACHYKSVLRVLRGHHIRVNEG 60
Db 24 IVGGTTPESHVPYQVSLNSGVHFCGSLINDQVWVSAACHYKSRIVQLRGEHNIVLEG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNIMLIKLTXPATLNOYVHAVALPTFCADATMCTVSG 120
Db 84 DEQFINAAKIIRPNYSMTLNDIMLIKUSPVKLNRVAVPALSPACAPAGTQCLISG 143
Qy 121 WGNMTSS--VADGDKLXKLPLTSLHADCAANSYGMITQSMFCAGYLEGGKDCQDSGG 179
Db 144 WGNLTSLNGVNPDLQCVDAFVPLSQADCEAA--PGEITSSMFCVGLSGKDCQDSGG 202
Qy 180 PWCNGLQGVGWSNGYGCABNDXPGYAKVXVLSGWRDTMA 223
Db 203 PWCNGQLQ--GIVSWGVCALPNDPFGVTKVNFVGMITQTTA 244
Query Match 62.8%; Score 727; DB 1; Length 247;
Best Local Similarity 61.6%; Pred. No. 9e-64;
Matches 138; Conservative 31; Mismatches 51; Indels 4; Gaps 3;

RESULT 10
B25528
Qy 1 IVGGYECTHSQAHQVSLNSGVKXGSLINXWVSAACHYKSVLRVLRGHHIRVNEG 60

```

```

trypsin (EC 3.4.21.4) precursor - mouse
C|Species: Mus musculus (house mouse)
C|Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C|Accession: B25528
R|Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
Nucleic Acids Res. 14, 8307-8310, 1986
A|Title: Sequence organization and transcriptional regulation of the mouse elastase II at
A|Reference number: A93446; MUID: 87066713; PMID:3641189
A|Accession: B25528
A|Molecule type: mRNA
A|Residues: 1-246 <STE>
A|Cross-references: GB:X04574; NID:954918; PIDN:CAA28243.1; PID:g54919
C|Superfamily: trypsin; trypsin homology
C|Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F|1-23/Domain: signal sequence #status predicted <SIG>
F|24-246/Product: trypsin #status predicted <MAT>
F|24-239/Domain: trypsin homology <TRY>
F|30-160,48-64,132-233,139-206,171-185/Diulfide bonds: #status predicted
F|63,107,200/Active site: His, Asp, Ser #status predicted
F|75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 63.0%; Score 729; DB 2; Length 246;
Best Local Similarity 61.2%; Pred. No. 5.7e-64;
Matches 137; Conservative 27; Mismatches 56; Indels 4; Gaps 3;

Qy 1 IVGGYECTHSQAHQVSLNSGVKXGSLINXWVSAACHYKSVLRVLRGHHIRVNEG 60
Db 24 IVGGYTCRESVPYQVSLNSGVHFCGSLINDQVWVSAACHYKTRIVQLRGEHNIVLEG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNIMLIKLTXPATLNOYVHAVALPTFCADATMCTVSG 120
Db 84 NEQFVDSAKIIRPNYSMTLNDIMLIKUSPVLNRVAVPALSPACAPAGTQCLISG 143
Qy 121 WGNMTSS--VADGDKLXKLPLTSLHADCAANSYGMITQSMFCAGYLEGGKDCQDSGG 179
Db 144 WGNLTSLNGVNPDLQCVDAFVPLSQADCEAA--PDDITNMICVGLSGKDCQDSGG 202
Qy 180 PWCNGLQGVGWSNGYGCABNDXPGYAKVXVLSGWRDTMA 223
Db 203 PWCNGELQ--GIVSWGVCALPNDPFGVTKVNFVGMITQTTA 244
RESULT 11
TRDG
trypsin (EC 3.4.21.4) precursor, anionic - dog
N|Alternate names: cationic trypsinogen
C|Species: Canis lupus familiaris (dog)
C|Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C|Accession: A26273
R|Pinsky, S.D.; LaForge, K.S.; Scheele, G.
Mol. Cell. Biol. 5, 2669-2676, 1985
A|Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA seque
A|Reference number: A26273; MUID: 86284628; PMID:3841794
A|Accession: A26273
A|Molecule type: mRNA
A|Residues: 1-247 <PIN>
A|Cross-references: GB:M11589; NID:g164094; PIDN:AAA30899.1; PID:g164095
C|Superfamily: trypsin; trypsin homology
C|Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F|1-15/Domain: signal sequence #status predicted <SIG>
F|16-23/Domain: activation peptide #status predicted <APT>
F|24-247/Product: trypsin, anionic #status predicted <EN2>
F|24-239/Domain: trypsin homology <TRY>
F|30-160,48-64,132-233,139-206,171-185/Diulfide bonds: #status predicted
F|63,107,200/Active site: His, Asp, Ser #status predicted
F|75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 62.8%; Score 727; DB 1; Length 247;
Best Local Similarity 61.6%; Pred. No. 9e-64;
Matches 138; Conservative 31; Mismatches 51; Indels 4; Gaps 3;

Qy 1 IVGGYECTHSQAHQVSLNSGVKXGSLINXWVSAACHYKSVLRVLRGHHIRVNEG 60

```

```

Db 24 IVGGVCTEENSVPYQVSLNAGHYFCGSLISDQWVSAACHYKSRIQVRLGEHINVDLWG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 120
Db 84 NEQFINAKVIRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 143
Qy 121 WGNWTSSVXD-GDKLOXILPILSHADCNISYGMKTSQMFPCAGYLEGKDCSQGDSGG 179
Db 144 WGNLTSSGNYEPQLQCLDAPLTPQACEASY-PQITINMVCAGFLEGKDCSQGDSGG 202
Qy 180 PIVCNGVLAGVWSNGVGCARERDPCGVYAKYVLSGWVDTMA 223
Db 203 PIVCNGELO--GIVSMGYCAQKNGKGVYTVKCNVDMVDMIOSTIA 244

RESULT 12
TSTR2
trypsin (EC 3.4.21.4) II precursor - rat
N:Alternate names: trypsinogen II
C:Species: Rattus norvegicus (Norway rat)
C>Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 18-Jul-1997
C:Accession: A22657; A00949
C:Suprafamily: trypsin; protein digestion; serine proteinase
A:Title: Structure of the active site of rat pancreatic trypsin genes.
A:Reference number: A22657; PMID:85054860; PMID:6094547
A:Accession: A22657
A:Molecule type: DNA
A:Residues: 1-246 <CRA>
R:MacDonald, R.J.; Swift, S.J.; Stary, G.H.
J. Biol. Chem. 257, 9724-9732, 1982
A:Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
A:Reference number: A00948; PMID:82265624; PMID:6896710
A:Accession: A00949
A:Molecule type: mRNA
A:Residues: 9-246 <MAC>
C:Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I mRNA.
C:Genetics: 14/1; 67/2
A:Suprafamily: trypsin; trypsin homology
C:Keywords: hydrolase, pancreas, protein digestion; serine proteinase; zymogen
F:1-15/Domain: signal peptide; status predicted <SIG>
F:16-23/Domain: activation peptide; status predicted <APT>
F:24-246/Product: trypsin II #status predicted <ENR>
F:24-239/Domain: trypsin homology <TRY>
F:30-160, 48-64, 132-233, 139-206, 171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 62.7%; Score 726; DB 1; Length 246;
Best Local Similarity 62.1%; Pred. No. 1.1e-63;
Matches 139; Conservative 24; Mismatches 57; Indels 4; Gaps 3;

Qy 1 IVGGVCTKHSQAHQVSLNAGHYFCGSLINXXVWSAACHYKSVLRVLRGHHIRVNEG 60
Db 24 IVGGVCTQENSVPYQVSLNAGHYFCGSLINQWVWSAACHYKSRIQVRLGEHINVDLWG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 120
Db 84 DEQFINAKVIRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 143
Qy 121 WGNWTSSVXD-GDKLOXILPILSHADCNISYGMKTSQMFPCAGYLEGKDCSQGDSGG 179
Db 144 WGNLTSSGNYEPQLQCLDAPLTPQACEASY-PQITINMVCAGFLEGKDCSQGDSGG 202
Qy 180 PIVCNGVLAGVWSNGVGCARERDPCGVYAKYVLSGWVDTMA 223
Db 203 PIVCNGELO--GIVSMGYCALPNQKNGKGVYTVKCNVDMVDMIOSTIA 244

RESULT 13
S13813
trypsin (EC 3.4.21.4) - bovine

```

```

C:Species: Bos primigenius taurus (cattle)
C>Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C:Accession: S13813
R:Le Huec, I.; Wicker, C.; Guilloteau, P.; Toulle, R.; Puigserver, A.
Eur. J. Biochem. 193, 767-773, 1990
A:Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsin.
A:Reference number: S13813; PMID:91065383; PMID:1701147
A:Accession: S13813
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-247 <HUE>
C:Cross-reference: EMBL:X54703; NID:9829; PIDN:CAA38513.1; PID:g930
C:Suprafamily: trypsin; trypsin homology
C:Keywords: hydrolase, pancreas, protein digestion; serine proteinase
F:1-15/Domain: signal peptide; status predicted <SIG>
F:24-239/Domain: trypsin homology <TRY>
F:63,107,200/Active site: His, Asp, Ser #status predicted

Query Match 62.3%; Score 721; DB 2; Length 247;
Best Local Similarity 60.7%; Pred. No. 3.5e-63;
Matches 136; Conservative 30; Mismatches 54; Indels 4; Gaps 3;

Qy 1 IVGGVCTKHSQAHQVSLNAGHYFCGSLINXXVWSAACHYKSVLRVLRGHHIRVNEG 60
Db 24 IVGGVCTAENSVPYQVSLNAGHYFCGSLINQWVWSAACHYQVRLGEHINVDLWG 83
Qy 61 TEQXISSSVXRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 120
Db 84 GEQFDASKVIRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 143
Qy 121 WGNWTSSVXD-GDKLOXILPILSHADCNISYGMKTSQMFPCAGYLEGKDCSQGDSGG 179
Db 144 WGNLTSSGNYEPQLQCLDAPLTPQACEASY-PQITINMVCAGFLEGKDCSQGDSGG 202
Qy 180 PIVCNGVLAGVWSNGVGCARERDPCGVYAKYVLSGWVDTMA 223
Db 203 PIVCNGELO--GIVSMGYCAQKNGKGVYTVKCNVDMVDMIOSTIA 244

RESULT 14
A27547
trypsin (EC 3.4.21.4) precursor, cationic - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1998 #sequence_revision 30-Jun-1998 #text_change 22-Jun-1999
C:Accession: A27547
C:Suprafamily: trypsin; trypsin homology
C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F:25-240/Domain: trypsin homology <TRY>
F:31-161,49-65,133-234,140-207,172-186/Disulfide bonds: #status predicted
F:64,108,201/Active site: His, Asp, Ser #status predicted
F:76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 62.1%; Score 719; DB 2; Length 247;
Best Local Similarity 61.2%; Pred. No. 5.5e-63;
Matches 137; Conservative 26; Mismatches 57; Indels 4; Gaps 3;

Qy 1 IVGGVCTKHSQAHQVSLNAGHYFCGSLINXXVWSAACHYKSVLRVLRGHHIRVNEG 60
Db 25 IVGGVCTQKNSLPYQVSLNAGHYFCGSLINQWVWSAACHYKSRIQVRLGEHINVDLWG 84
Qy 61 TEQXISSSVXRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 120
Db 85 GEQFDAAKIRHPNYSYNIKNDIMLIXTPATLNQYHVAALPTECAADATMCTVSG 144
Qy 121 WGNWTSSVXD-GDKLOXILPILSHADCNISYGMKTSQMFPCAGYLEGKDCSQGDSGG 179
Db 144 WGNLTSSGNYEPQLQCLDAPLTPQACEASY-PQITINMVCAGFLEGKDCSQGDSGG 202

```

Db 145 WGNLSSGTYPSLLOCLDAPVLSDSCKSSY-PGKITSNMPCLGFLGKDSQGDGG 203

QY 180 PVV CNGVLQGVGVVSWGYGCAERDXPGVYAKVXVLSGWVRDTMA 223

204 PWVCNGOLO--GVVSWGYGCAOKGKPGVYTKVCNVVNWIOOTVA 245

RESULT 15

RESOL
A35871

A33871
trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog

C:Species: *Xenopus laevis* (African clawed frog)

C:\Spectra\Xenopus laevis (Allison Crawd 1109)
C:\Date: 09-Nov-1990 #sequence revision 09-Nov-1990 #text change 04-Mar-2000

C:\DATE: 09-NOV-1990 #SEQUENCE
C:\ACCESSION: A35871: S12117

C;Accession: A35871; SL2:
P:Shi Y B : Brown D D

R;SHL, Y.B.; BROWN, D.D.
Genes Dev 4 1107-1113 1990

Genes Dev. 4, 1107-1113, 1990

A; Title: Developmental and thyroid hormone-dependent regulation of the expression of the *hsp70* gene in the developing rat brain. Reference number: 325971. MID: 91007355. BMID: 22103372

A;Reference number:
A;Accession: 225071

A;Accession: A35871
A;Status: preliminary

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-243 <SHI>

A;Cross-references: EMBL:X53458; NID:g65

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; protein digestion; serine proteinase

F;1-15/Domain: signal sequence #status predicted <SIG>

F;16-20/Domain: activation peptide #stat

F:21-236/Domain: trypsin homology <TRY>

```
F;26-243/Product: trypsin I #status predicted <MAT>
```

P; 27-157, 45-61, 129-230, 136-203, 168-182/Disulfide bonds: #

Query Match 62.0% Score 718: DB 2: Length 243;

Best Local Similarity 61.2%; pred. No. 6.7e-63;

BEBC LOCAL SIMILARITY	Q1:29	FREQ: NO. 0:76-85
Matches 137	Conservative 31	Mismatches 52
	Indels 4	Gaps 3

1 IVGGYECTXHSAHOVSLNSGYHXC GGSLINXXWTVSAAHCYKSVLRVRLGEHHIRVNEG 60

[illegible]

8 21 TGGATCAKSSVPIVSENNGIHF CGSIIINQWVSAATCINASTQVRKGEHNTALSEG 89

61 TEOYISSESVYBHPNVSSVNTYNDIM.IKI.TYRATI.NOVVHAVA.I.PTECAADATMCTVSG 120

[illegible]

b 81 TEQFISSSKVIRHSGYNSYTLNDIMLIKLSPPASLNAAVNTVPLPSGCSAAGTSCLSIG 140

121 WGNMSSVXD-GDKLQXLPLILSHADCANSYGPGMITQSMFCAGYLEGGKDS CQGDSSG 179

[illegible]

Search completed: February 12, 2003, 10:28:56

Job time : 57.4179 sec

OM protein - protein search, using sw model
 Run on: February 12, 2003, 10:04:45 ; Search time 30.2239 Seconds
 (without alignments)
 308.768 Million cell updates/sec

GenCore version 5.1.3
 Copyright (c) 1993 - 2003 CompuGen Ltd.

Perfect score: 1158

Sequence: 1 IVGGYECTMHSQAHQSLNS.....GVYAKYKVLSCWRDTWAKY 225

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086.5	93.8	241	1 TRV1_GADMO	P16049 gadus morhua
2	1085.5	93.7	241	1 TRVX_GADMO	P31041 gadus morhua
3	901.5	77.8	242	1 TRV1_SALSA	P15072 salmo salar
4	893.5	77.2	231	1 TRV2_SALSA	P15072 salmo salar
5	758.5	65.5	229	1 TRV2_SQUAC	P35033 salmo salar
6	756.5	65.3	238	1 TRV3_SALSA	P35033 salmo salar
7	755.5	65.2	246	1 TRV1_RAT	P00762 rattus norv
8	735.5	63.5	244	1 TRV2_XENLA	P70059 xenopus lae
9	729.5	63.0	246	1 TRV2_MOUSE	P07146 mus musculus
10	727.5	62.8	247	1 TRV2_CANFA	P06872 canis fami
11	726.5	62.7	246	1 TRV2_RAT	P00763 rattus norv
12	721.5	62.3	247	1 TRV3_RAT	Q23463 bos taurus
13	717.5	62.3	247	1 TRV1_XENLA	P08436 rattus norv
14	716.5	62.3	247	1 TRV2_BOVIN	P19799 xenopus lae
15	710.5	61.3	243	1 TRV2_BOVIN	P07478 homo sapien
16	707.5	61.1	231	1 TRV2_CANFA	P00761 canis fami
17	706.5	61.0	246	1 TRV1_CANFA	P06871 canis fami
18	702.5	60.6	247	1 TRV1_HUMAN	P04770 homo sapien
19	701.5	60.5	243	1 TRV1_BOVIN	P06297 bos taurus
20	697.5	60.2	248	1 TRV1_CHICK	Q90628 gallus gall
21	693.5	59.8	248	1 TRV2_CHICK	P35030 homo sapien
22	691.5	59.7	304	1 TRV4_HUMAN	Q90629 gallus gall
23	674.5	58.2	248	1 TRV3_CHICK	P15951 homo sapien
24	665.5	57.4	247	1 TRV4_HUMAN	P12788 rattus norv
25	650.5	56.2	247	1 TRV4_RAT	P32842 rattus norv
26	617.5	53.3	246	1 TRV4_RAT	P32841 rattus norv
27	616.5	53.2	246	1 TRV4_RAT	P35034 pleuronecte
28	529.5	45.7	250	1 TRYP_PLEPL	Q9H255 pleuronecte
29	471.5	40.7	256	1 KLAF_HUMAN	P35037 anopheles g
30	471.5	40.7	256	1 TRV3_ANGA	P35035 anopheles g
31	467.5	40.4	277	1 TRV2_ANGA	P35035 anopheles g
32	467.5	39.8	274	1 TRV1_ANGA	Q9P093 anopheles
33	460.5	39.7	251	1 KLX2_HUMAN	Q9P093 anopheles

34	458	39.6	267	1 TRV7_ANGA	P35041 anopheles g
35	451.5	38.9	250	1 KLX2_HUMAN	Q9ubx7 homo sapien
36	449.5	38.7	263	1 CTRA_GADMO	P47796 gadus morhu
37	448	38.6	261	1 KLX6_HUMAN	Q92876 gadus morhu
38	447.5	38.6	261	1 KLX6_MOUSE	P00756 mus musculu
39	446.5	38.6	261	1 KLX9_MOUSE	P15949 mus musculu
40	441.5	38.1	261	1 KLX2_HUMAN	P00151 gadus morhu
41	433.5	37.4	245	1 CTRB_GADMO	P00752 gadus morhu
42	433.5	37.4	246	1 KLK2_PIG	Q9uk73 homo sapien
43	432.5	37.3	277	1 TRYP_ASTEL	P00765 astacus flu
44	431	37.2	237	1 TRYP_HUMAN	P49852 homo sapien
45	431	37.2	253	1 KLK7_HUMAN	

ALIGNMENTS

RESULT 1	
TRV1_GADMO	
ID	TRV1_GADMO STANDARD;
AC	P16049; 091040; 092156;
DT	01-APR-1990 (Rel. 14, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Trypsin I precursor (EC 3.4.21.4).
OS	Gadus morhua (Atlantic cod).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX	NCBI_TaxID=8049;
RN	[1]
RC	SEQUENCE FROM N.A.
RC	TISSUE=Pyloric caeca;
RC	MEDLINE=94039130; PubMed=8221632;
RC	Gadumad02134; PubMed=8221632;
RA	Gadumad02134; PubMed=8221632;
RA	Eakin A.E., Craik C.S., Oskarsson S., Bjarnason J.B.,
RT	"Isolation and characterization of cDNAs from Atlantic cod encoding
RT	two different forms of trypsinogen."
RT	Eur. J. Biochem. 217:1091-1097(1993).
RL	[2]
RP	SEQUENCE OF 20-58.
RC	TISSUE=Pyloric caeca;
RC	MEDLINE=94039130; PubMed=2707266;
RA	Argelsson B., Fox J.W., Bjarnason J.B.;
RT	"Purification and characterization of trypsin from the poikilotherm
RT	Gadus morhua."
RL	Eur. J. Biochem. 180:85-94(1989).
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Arg- Xaa, Lys- Xaa.
CC	-1- SUBCELLULAR LOCATION: Extracellular.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is not modified and
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.ebi.ac.uk/ebis/announce/
CC	or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; X76886; CAA54214.1; -
DR	PIR; S03570; S03570.
DR	RSSP; P00763; IDPO.
DR	MEROPS; S01151; -
DR	InterPro; IPR001314; Chymotrypsin.
DR	Trypsin; P0001234; Ser. protease_Try.
DR	PIR; P00069; CHYMOTRYPSIN.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	SMART; SM00020; TRYPSPC1.
DR	PROSITE; PS02404; TRYPSPIN_DOM_1.
DR	PROSITE; PS00134; TRYPSPIN_HIS_1.
DR	PROSITE; PS00135; TRYPSPIN_SER_1.
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW	Multigene family.

FT SIGNAL 1 13 POTENTIAL.
FT PROPEP 14 19 ACTIVATION PEPTIDE.
FT CHAIN 20 241 TRYPSIN I.
FT ACT SITE 59 59 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 103 103 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 26 155 BY SIMILARITY.
FT DISULFID 44 60 BY SIMILARITY.
FT DISULFID 128 228 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 166 180 BY SIMILARITY.
FT DISULFID 191 215 BY SIMILARITY.
FT SITE 189 189 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
FT CONFLICT 25 28 TK -> Q (IN REF. 2).
FT CONFLICT 31 31 F -> Y (IN REF. 2).
FT CONFLICT 43 43 VSQD -> Y (IN REF. 2).
FT CONFLICT 49 52
SQ SEQUENCE 241 AA; 25941 MW; 448C9A0106AD1A68 CRC64;

Query Match 93.8%; Score 1086.5; DB 1; Length 241;
Best Local Similarity 91.6%; Pred. No. 1.9e-104;
Matches 206; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

Oy 1 IVGVGYCTKSHQAHQVSLNSGVHKGCGSLINXWVSAACHYKSLVRLGEGHHRVNEG 60
Db 20 IVGVGYCTKSHQAHQVSLNSGVHKGCGSLVSKDWVSAACHYKSLVRLGEGHHRVNEG 79

Oy 61 TEQXISSSSVRHPNYSYNIKNDIMLIKLTXPATLNQYVHAVALPTECAADATMCTVSG 120
Db 80 TEQVTSSSSVIRHPNYSYNIKNDIMLIKLTXPATLNQYVHAVALPTECAADATMCTVSG 139

Oy 121 WGMTSSVVDGDKLQXKLPLSHADCANSGPGMITQSMFCAGYLEGGKSCQSGSGP 180
Db 140 WGMTSSVVDGDKLQXKLPLSHADCANSGPGMITQSMFCAGYLEGGKSCQSGSGP 198

Oy 181 VVCGVGLQGVVSNVSGYGCARDPFGYAKVXVLSGWRDTWAXY 225
Db 199 VVCGVGLQ--GVVSNVSGYGCARDPFGYAKVXVLSGWRDTWAXY 241

RESULT 2
TRYX GADMO STANDARD; PRT; 241 AA.
AC Q91041.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin I precursor (EC 3.4.21.4).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN NCBI_TaxID=8049;
RP SEQUENCE FROM N.A.
RC TISSUE=Pyloric caeca;
RX MEDLINE=94039130; PubMed=8223632;
RA Gudmundsdottir A., Gudmundsdottir E., Oskarsson S., Bjarnason J.B.,
BA Eakin R.E., Craik C.S.
RT "Isolation and characterization of cDNAs from Atlantic cod encoding two different forms of trypsinogen."
RL Eur. J. Biochem. 217:1091-1097(1993).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to license@isb-sib.ch).

CC EMBL; X76887; CAA54215.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.151; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00400; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 13 POTENTIAL.
FT PROPEP 14 19 ACTIVATION PEPTIDE.
FT CHAIN 20 241 TRYPSIN X.
FT ACT SITE 59 59 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 103 103 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 26 155 BY SIMILARITY.
FT DISULFID 44 60 BY SIMILARITY.
FT DISULFID 128 228 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 166 180 BY SIMILARITY.
FT DISULFID 191 215 BY SIMILARITY.
FT SITE 189 189 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 241 AA; 25976 MW; 853D7C68C9FDD7 CRC64;

Query Match 93.7%; Score 1085.5; DB 1; Length 241;
Best Local Similarity 91.6%; Pred. No. 2.4e-104;
Matches 206; Conservative 2; Mismatches 14; Indels 3; Gaps 2;

Oy 1 IVGVGYCTKSHQAHQVSLNSGVHKGCGSLINXWVSAACHYKSLVRLGEGHHRVNEG 60
Db 20 IVGVGYCTKSHQAHQVSLNSGVHKGCGSLVSKDWVSAACHYKSLVRLGEGHHRVNEG 79

Oy 61 TEQXISSSSVRHPNYSYNIKNDIMLIKLTXPATLNQYVHAVALPTECAADATMCTVSG 120
Db 80 TEQVTSSSSVIRHPNYSYNIKNDIMLIKLTXPATLNQYVHAVALPTECAADATMCTVSG 139

Oy 121 WGMTSSVVDGDKLQXKLPLSHADCANSGPGMITQSMFCAGYLEGGKSCQSGSGP 180
Db 140 WGMTSSVVDGDKLQXKLPLSHADCANSGPGMITQSMFCAGYLEGGKSCQSGSGP 198

Oy 181 VVCGVGLQGVVSNVSGYGCARDPFGYAKVXVLSGWRDTWAXY 225
Db 199 VVCGVGLQ--GVVSNVSGYGCARDPFGYAKVXVLSGWRDTWAXY 241

RESULT 3
TRYI SALSIA STANDARD; PRT; 242 AA.
AC P35031;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin I precursor (EC 3.4.21.4).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN NCBI_TaxID=8030;
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=96035908; PubMed=755623;
RA Male R., Lorenz J.B., Small A.O., Torrisen K.R.;
RT "Molecular cloning and characterization of anionic and cationic variants of trypsin from Atlantic salmon."
RL Eur. J. Biochem. 232:677-685(1995).
CC
CC X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS).

RA Smalas A.O., Hordvik A.;
 PT "Structure determination and refinement of benzamidine-inhibited
 PT trypsin from the North Atlantic salmon (Salmo salar) at 1.82-A
 RL resolution";
 RL Acta Crystallogr. D 49:318-330(1993).
 [3].
 RP X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS).
 RA MEDLINE=95148588; PubMed=7846025.
 RA Smalas A.O., Heimstad E.S., Hordvik A., Willassen N.P., Male R.;
 RT "Cold adaptation of enzymes: structural comparison between salmon and
 RT bovine trypsins";
 RL Proteins 20:149-166(1994).
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- SUBCELLULAR LOCATION: Extracellular.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X70075; CAA49680.1; -;
 DR EMBL; X70071; CAA49676.1; -;
 DR EMBL; X70072; CAA49677.1; -;
 DR PIR; S31775; S31775.
 DR PIR; S31776; S31776.
 DR PIR; S31777; S31777.
 DR PDB; 2TBS; 30-APR-94.
 DR PDB; 1BIT; 01-NOV-94.
 DR MEROPS; S01.151; -;
 DR InterPro; IPRO00134; Ser_protease_Try.
 DR Pfam; PF000489; Trypsin_MOTIFPSIN.
 DR PRINTS; PR00722; Trypsin_MOTIFPSIN.
 DR SMART; SM00020; Tryp_Spc_1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family; 3D-structure.
 KM
 FT SIGNAL 1 15 POTENTIAL.
 FT PROPEP 16 20 ACTIVATION PEPTIDE.
 FT CHAIN 21 242 TRYPSIN I.
 FT ACT_SITE 60 60 CHARGE RELAY SYSTEM.
 FT ACT_SITE 104 104 CHARGE RELAY SYSTEM.
 FT ACT_SITE 196 196 CHARGE RELAY SYSTEM.
 FT DISULFID 27 156 CHARGE RELAY SYSTEM.
 FT DISULFID 145 51
 FT DISULFID 145 229
 FT DISULFID 136 181
 FT DISULFID 167 181
 FT DISULFID 192 216
 FT SITE 190 190
 FT SITE 33 33 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 FT VARIANT 33 33 T -> A (IN TRYPSIN I)/TD.
 FT SEQUENCE 242 AA; 25958 WW; 43F5642498067EEA CRC64;
 Query Match 77.8%; Score 901.5; DB 1; Length 242;
 Best Local Similarity 75.1%; Pred. No. 1.9e-85;
 Matches 169; Conservative 15; Mismatches 38; Indels 3; Gaps 2;
 QY 1 IGVGVCTKHSHQHVNSGVHCGSLINXWVVAACHYKSVLRLGEHHRVNEG 60
 DB 2 IGVGVCTKHSHQHVNSGVHCGSLINXWVVAACHYKSVLRLGEHHRVNEG 60
 QY 61 TEQVTSSTSSVSEHNSYKYNINIMLKITYPATLQVYHVAFLPTCEAATMTVSG 120
 DB 81 SEQVTSSTSSVSEHNSYKYNINIMLKITYPATLQVYHVAFLPTCEAATMTVSG 140
 QY 121 WGMTNMTSVDKDLKXLPILSHLGSNKPATLVQVAPLTFSCAPAGTMTVSG 180
 DB 121 WGMTNMTSVDKDLKXLPILSHLGSNKPATLVQVAPLTFSCAPAGTMTVSG 180

DB 141 WGMTNMTSVDKDLKXLPILSHLGSNKPATLVQVAPLTFSCAPAGTMTVSG 199
 QY 181 WGVNVLQGVNSGVHCGSLINXWVVAACHYKSVLRLGEHHRVNEG 225
 DB 200 WGVNVLQGVNSGVHCGSLINXWVVAACHYKSVLRLGEHHRVNEG 242
 RESULT 4
 TRV2_SALSA
 ID TRV2_SALSA STANDARD; PRT; 231 AA.
 AC P35032;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 OS Trypsin II Precursor (EC 3.4.21.4) (Fragment).
 CS Trypsin II Precursor (EC 3.4.21.4) (Fragment).
 CC Eukaryotic; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Acinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OK NCBI_TaxID=8030.
 RN [1] _SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=96035908; PubMed=7556223;
 RA Male R., Lorens J.B., Smals A.O., Torrisen K.R.;
 RT "Molecular cloning and characterization of anionic and cationic
 RT variants of trypsin from Atlantic salmon";
 RL Eur. J. Biochem. 232:677-685(1995).
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- SUBCELLULAR LOCATION: Extracellular.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X70073; CAA49678.1; -;
 DR EMBL; X70073; CAA49678.1; -;
 DR PIR; S31778; S31778.
 DR HSP; P35031; IBIT.
 DR MEROPS; S01.258; -;
 DR InterPro; IPRO001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp_Spc_1.
 DR PROSITE; PS00140; TRYPSIN_DOM; 1.
 DR PROSITE; PS00140; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family.
 KM
 FT NON_TER 1 1
 FT SIGNAL <1 4 POTENTIAL.
 FT PROPEP 5 9 ACTIVATION PEPTIDE.
 FT CHAIN 10 231 TRYPSIN II.
 FT ACT_SITE 49 49 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 93 93 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 185 185 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 16 145 BY SIMILARITY.
 FT DISULFID 34 50 BY SIMILARITY.
 FT DISULFID 118 218 BY SIMILARITY.
 FT DISULFID 125 191 BY SIMILARITY.
 FT DISULFID 136 200 BY SIMILARITY.
 FT DISULFID 180 200 BY SIMILARITY.
 FT SITE 179 179 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 231 AA; 24823 WW; C54CAFE74FAE18 CRC64;
 Query Match 77.2%; Score 893.5; DB 1; Length 231;
 Best Local Similarity 74.7%; Pred. No. 1.2e-84;
 Matches 168; Conservative 16; Mismatches 38; Indels 3; Gaps 2;
 QY 1 IGVGVCTKHSHQHVNSGVHCGSLINXWVVAACHYKSVLRLGEHHRVNEG 60

```

DR PIR; A00950; TRDES.
DR HSSP; P00763; IDPO.
DR MEROPS; S01151; -.
DR InterPro; IP0001314; Chymotrypsin.
DR InterPro; IP0001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
KW PROPEP
FT CHAIN 1
FT ACT_SITE 49
FT ACT_SITE 91
FT ACT_SITE 183
FT DISULFID 14
FT DISULFID 32
FT DISULFID 116
FT DISULFID 123
FT DISULFID 146
FT DISULFID 154
FT DISULFID 179
FT SITE 177
FT VARIANT 109
FT SEQUENCE 229 AA; 24591 MW; E83B83C5AD72FCE4 CRC64;
Query Match 65.5%; Score 758.5; DB 1; Length 229;
Best Local Similarity 62.8%; Pred. No. 96-71;
Matches 140; Conservative 23; Mismatches 57; Indels 3; Gaps

```

CC	EMBL	X70074	CAA49679.1	-	
CC	PIR	S31779	S31779		
DR	HSSP	P00763	IDPO		
DR	MEROPS	S01151	-		
DR	InterPro	IPR001254	Ser_protease_Try		
DR	Pfam	PF00089	trypsin_1		
DR	SMART	SM00020	Tryp_Spc_1		
DR	PROSITE	PS02040	TRYPSIN_DOM_1		
DR	PROSITE	PS00134	TRYPSIN_HIS_1		
DR	PROSITE	PS00135	TRYPSIN_SER_1		
KW	Hydrolase		Serine protease; Digestion; Pancreas; Zymogen; Signal;		
KW	Multigene family.				
KW	NON_TER		1		
FT	SIGNAL		<1	15	POTENTIAL.
FT	PROPEP		8	15	ACTIVATION PEPTIDE.
FT	CHAIN		56	235	TRYPSIN III.
FT	ACT_SITE		56	58	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE		99	99	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE		192	192	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID		22	152	BY SIMILARITY.
FT	DISULFID		40	56	BY SIMILARITY.
FT	DISULFID		124	235	BY SIMILARITY.

FT DISULFID 131 198 BY SIMILARITY.
 FT DISULFID 163 177 BY SIMILARITY.
 FT DISULFID 188 212 BY SIMILARITY.
 FT SEQ 186 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 236 AA; A679980E8393023 CRC64;
 Query Match 65.3%; Score 756; DB 1; Length 236;
 Best Local Similarity 63.8%; Fragment No. 1.7e-70;
 Matches 143; Conservative 26; Mismatches 51; Indels 4; Gaps 3;
 QY 1 IVGVGYECTHQHQVSLNCGXGSLNXXVYSAAHCYKSLVPLVCEHHTRNNG 60
 DB 16 IVGVGYEGRKNSAYQASLQGVHFCGSLISSTWVSAHCYKSRIVQLSEIANVNS 75
 QY 61 TEQVSSSSVXRPNTSSVYNXDMLIKLTXPATLNQVVAVALPTECAADATMCTVSG 120
 DB 76 TEQVSDVSYVHPHSVNSNDLMLIKLSKPASLNSVYVSTVLPSSCASSOTCLVSG 135
 QY 121 NQNTMSVYD-GDKLKLPLISLHSDANSYFGQMITQSMFCAGYLEGGKSCQGSGG 179
 DB 136 NQNLSSGSSNYPDLKDLPLISSSSCSNAY-PCQITSNMFCAGMEGGKSCQGS 194
 QY 180 PIVGVGVQGVVSGVSGCAEDYGVYAKVXVLSGWDPDMA 223
 DB 195 PIVGVGVQGVVSGVSGCAEDYGVYAKVXVLSGWDPDMA 223
 DB 195 PIVGVGVQGVVSGVSGCAEDYGVYAKVXVLSGWDPDMA 223
 RESULT 7
 TRV1 RAT
 ID TRV1 RAT STANDARD; PRT; 246 AA.
 AC P00762;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 25-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin I, anionic precursor (EC 3.4.21.4) (Pretrypsinogen I).
 GN Trypsin I, anionic precursor (EC 3.4.21.4) (Pretrypsinogen I).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1] _SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=82265624; PubMed=6896710;
 RA McDonald R.J., Stary S.J., Swift G.H.;
 RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
 RT sequences of the cloned cDNAs."
 RL J. Biol. Chem. 257:9724-9732(1982).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=850594880; PubMed=6094547;
 RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
 RA Rutter W.J., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
 RT "Structure of two related rat pancreatic trypsin genes."
 RL J. Biol. Chem. 259:14255-14264(1984).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=87292123; PubMed=3112942;
 RA Sprang S., Standing T., Fletcher R.J., Stroud R.M., Finer-Moore J.,
 RA Xuong N.-H., Hamlin R., Rutter W.J., Craik C.S.;
 RT "The three-dimensional structure of An102 mutant of trypsin: role of
 RT Asp102 in serine protease catalysis."
 RL Science 237:905-909(1987).
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- BIOSYNTHESIS: THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR
 CC FORM OF TRYPSIN PRODUCED BY THE ADULT PANCREAS.
 CC -!- SIMILARITY: BELONGS TO PAPIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or email to license@isb-sib.ch).
 CC -----
 DR EMBL; V01273; CAA24580.1;
 DR EMBL; J00778; AAA98518.1;
 DR EMBL; A00948; TRST1.
 DR PDB; 1TRM; 15-JUL-93.
 DR PDB; 2TRM; 16-JUL-88.
 DR PDB; 1BRA; 30-APR-94.
 DR PDB; 1BRB; 31-JUL-94.
 DR PDB; 1BRC; 31-MAY-94.
 DR MEROPS; S01.094;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00093; Trypsin1.
 DR PRINTS; SM00072; Chymotrypsin.
 DR SMART; SM00072; Chymotrypsin.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PSS0135; TRYPSIN_SER; 1.
 KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family; 3D-structure.
 FT SIGNAL 1 15
 FT PROPEP 16 23 ACTIVATION PEPTIDE.
 FT CHAIN 24 246
 FT TRYPSIN I, ANIONIC.
 FT ACT_SITE 63 63 CHARGE RELAY SYSTEM.
 FT ACT_SITE 107 107 CHARGE RELAY SYSTEM.
 FT ACT_SITE 200 200 CHARGE RELAY SYSTEM.
 FT DISULFID 30 160
 FT DISULFID 48 64
 FT DISULFID 132 233
 FT DISULFID 171 186
 FT DISULFID 196 220
 FT SITE 194 194 REQUIRED FOR SPECIFICITY.
 FT STRAND 25 25
 FT STRAND 28 29
 FT TURN 32 33
 FT TURN 36 37
 FT STRAND 38 42
 FT STRAND 46 54
 FT TURN 55 56
 FT STRAND 57 60
 FT HELIX 62 64
 FT STRAND 70 73
 FT STRAND 77 77
 FT TURN 86 86
 FT TURN 88 88
 FT TURN 97 98
 FT STRAND 100 100
 FT TURN 101 104
 FT STRAND 105 105
 FT TURN 106 106
 FT STRAND 109 113
 FT STRAND 127 127
 FT TURN 135 136
 FT STRAND 138 143
 FT STRAND 157 157
 FT STRAND 159 165
 FT HELIX 168 174
 FT TURN 176 178
 FT TURN 181 182
 FT STRAND 183 186
 FT TURN 189 193
 FT STRAND 194 194
 FT TURN 197 198
 FT TURN 200 201
 FT STRAND 203 206
 FT TURN 207 208
 FT STRAND 209 216
 FT TURN 223 224
 FT STRAND 227 231

```
FT HELIX 232 234
FT TURN 236 244
FT TURN 245 246
SQ SEQUENCE 246 AA; 2959 MW; 6AFAADAD1943FB5 CRC64;

Query Match 65.28; Score 755; DB 1; Length 246;
Best Local Similarity 63.88; Pred. No. 2.2e-70;
Matches 143; Conservative 25; Mismatches 52; Indels 4; Gaps 3;

QY 1 IGVGECTHSQAHQVSLNSGYHCKGSLINXVTVVSAACHYKSVLRVLRGHHIRVNEG 60
DB 24 IGVGYTCPSHVPQVSLNSGYHCKGSLINDQWVSAACHYKSVLRVLRGHHIRVNEG 83
QY 61 TQXISSSSVXKHPNYSYXNDIMLKLTXPATLNQVVAVALPTECAADATMCTVSG 120
DB 84 DEQPINAAKIKHPNYSYXNDIMLKLTXPATLNQVVAVALPTECAADATMCTVSG 143
QY 121 WNTNMS-VXDGKQLXLPILSHADCANSGPGMITOSMFCAGLEGKDCSCGDSGG 179
DB 144 WNTLSNGWNPOLLQCDVAPVLSQADCEAAV-PGEITSMTCVGFLEGKDCSCGDSGG 202
QY 180 PIVCNQVGVGVSVGSGCAERDXPGVYAKVXVLSGWRDTMA 223
DB 203 PIVCNGQLQ--GVVSGYGCALPDNPGVYTKVCFVGMQDTIA 244

RESULT 8
TRY2_XENIA
ID TRY2_XENIA STANDARD; PRT; 244 AA.
AC P70059;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A. Gan L., Hood L.E.
RA Wang K., Jyic (1996) to the EMBL/GenBank/DBJ databases.
RL Submitted (1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC ENBL; U72330; AAB1274.1; --
CC DR HSPS; P00763; IDPO.
CC DR MEROPS; S01.258; --
CC
CC InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR001254; Ser protease_Try.
CC DR Pfam; P00089; trypsin; 1.
CC DR PRINTS; P00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; Tryp_SPC; 1.
CC DR PROSITE; PS0240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC DR Hydrolase; Serine protease; Digestion; Zymogen; Signal;
KW Multigene family.
KW SIGNAL 1 15 BY SIMILARITY.
FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 22 244 TRYPSIN.
FT ACT_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT

FT ACT_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 28 158 BY SIMILARITY.
FT DISULFID 46 62 BY SIMILARITY.
FT DISULFID 130 231 BY SIMILARITY.
FT DISULFID 137 204 BY SIMILARITY.
FT DISULFID 169 183 BY SIMILARITY.
FT DISULFID 194 218 BY SIMILARITY.
FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 244 AA; 26079 MW; C63F29CB3300B323 CRC64;

Query Match 63.58; Score 735; DB 1; Length 244;
Best Local Similarity 61.68; Pred. No. 2.5e-68;
Matches 138; Conservative 33; Mismatches 49; Indels 4; Gaps 3;

QY 1 IGVGECTHSQAHQVSLNSGYHCKGSLINXVTVVSAACHYKSVLRVLRGHHIRVNEG 60
DB 22 IGVGCTCAKNVAVFQVSLNSGYHCKGSLINQWVSAACHYKSVLRVLRGHHIRVNEG 81
QY 61 TQXISSSSVXKHPNYSYXNDIMLKLTXPATLNQVVAVALPTECAADATMCTVSG 120
DB 82 TQFIDQKQVIRHPNYSYXNDIMLKLTXPATLNQVVAVALPTECAADATMCTVSG 141
QY 121 WNTNMSVXD--GDKQLXLPILSHADCANSGPGMITOSMFCAGLEGKDCSCGDSGG 179
DB 142 WNTLSNGWNPOLLQCDVAPVLSQADCEAAV-PGEITSMTCVGFLEGKDCSCGDSGG 200
QY 180 PIVCNQVGVGVSVGSGCAERDXPGVYAKVXVLSGWRDTMA 223
DB 201 PIVCNGQLQ--GVVSGYGCALPDNPGVYTKVCFVGMQDTIA 242

RESULT 9
TRY2_MOUSE
ID TRY2_MOUSE STANDARD; PRT; 246 AA.
AC P07116;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DB Trypsin II, anionic precursor (EC 3.4.21.4) (Prettypinogen II).
GN TRY2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/J.
RX MEDLINE=87066713; PubMed=3641189;
RA Stevenson B.J., Hagenbuchle O., Wellauer P.K.;
RT "Sequence organisation and transcriptional regulation of the mouse
RT elastase II and trypsin genes."
RL Nucleic Acids Res. 14:8307-8350 (1986).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC ENBL; X04574; CAA28243.1; --
CC DR HSPS; P00763; IDPO.
CC DR MEROPS; S01.258; --
CC
CC InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR001254; Ser protease_Try.
CC DR Pfam; P00089; trypsin; 1.
CC DR PRINTS; P00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; Tryp_SPC; 1.
CC DR PROSITE; PS0240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC DR Hydrolase; Serine protease; Digestion; Zymogen; Signal;
KW Multigene family.
KW SIGNAL 1 15 BY SIMILARITY.
FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 22 244 TRYPSIN.
FT ACT_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT
```

```

DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_HIS; 1.
KW HistoLase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 16 23
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
SQ SEQUENCE 246 AA; 26203 MW; CF8F9YAC2D07AD CRC64;

Query Match
Best Local Similarity 63.0%; Score 729; DB 1; Length 246;
Matches 137; Conservative 27; Mismatches 56; Indels 4; Gaps 3;

QY 1 IVGGVCTCHSQAQVNSGYSYKCGSLINKXVYSANHCYKVLRLVGLGHHHNVGG 60
DB 24 IVGGVCTCHSQAQVNSGYSYKCGSLINKXVYSANHCYKVLRLVGLGHHHNVGG 60
QY 61 TEQXISSSVXKRPVNSYXNDIMLIKLTXPATLNOYVHVALPTCAADATVCTVSG 120
DB 84 NEQVDSAKIIRHPNYSNWTLDNIMLIKASPVTLNARVASVPLPSSCAPAGTCQLISG 143
QY 121 MNTNMSVXQDGLQXLMPLSHADCANSGYGMITQSMFCAGYLEGGKSCGDSGG 179
DB 144 MNTLSGNNVDPQLQVDAPVLPDQACASY-PGQITNNMTCVGLFEGKSCGDSGG 202
QY 180 PVQGVNQLQGVNSGVCGERDYPGYAKVYLSGWRDTPMA 223
DB 203 PVQGVNQLQ-GIVSWGTCAGQAKRPGVTKVCFVDMQSTIA 244

RESULT 10
TRY2 CANFA STANDARD; PRT; 247 AA.
AC P06872;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JUN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin, anionic precursor (EC 3.4.21.4).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxId=9615;
OX 11_TaxId=9615;
RP SEQUENCE FROM N.A.
RX MEDLINE=85284628; PubMed=3841794;
RA Pinsky S.D., Laforge K.S., Scheele G.;
RT "Differential regulation of trypsinogen mRNA translation: full-length
RT mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
RT in the dog pancreas."
RL Mol. Cell. Biol. 5:2669-2676 (1985).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. Non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).

```

```

CC
DR EMBL; M15499; AAA30899.1; -.
DR PIR; A24273; TRPC.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.258.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_HIS; 1.
KW HistoLase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Subtiligene family.
FT PROPEP 16 23
FT CHAIN 24 246
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
SQ SEQUENCE 247 AA; 26423 MW; 374E9D31D6DB8EAF CRC64;

Query Match
Best Local Similarity 62.8%; Score 727; DB 1; Length 247;
Matches 138; Conservative 31; Mismatches 51; Indels 4; Gaps 3;

QY 1 IVGGVCTCHSQAQVNSGYSYKCGSLINKXVYSANHCYKVLRLVGLGHHHNVGG 60
DB 24 IVGGVCTCHSQAQVNSGYSYKCGSLINKXVYSANHCYKVLRLVGLGHHHNVGG 60
QY 61 TEQXISSSVXKRPVNSYXNDIMLIKLTXPATLNOYVHVALPTCAADATVCTVSG 120
DB 84 NEQVDSAKIIRHPNYSNWTLDNIMLIKASPVTLNARVASVPLPSSCAPAGTCQLISG 143
QY 121 MNTNMSVXQDGLQXLMPLSHADCANSGYGMITQSMFCAGYLEGGKSCGDSGG 179
DB 144 MNTLSGNNVDPQLQVDAPVLPDQACASY-PGQITNNMTCVGLFEGKSCGDSGG 202
QY 180 PVQGVNQLQGVNSGVCGERDYPGYAKVYLSGWRDTPMA 223
DB 203 PVQGVNQLQ-GIVSWGTCAGQAKRPGVTKVCFVDMQSTIA 244

RESULT 11
TRY2 RAT STANDARD; PRT; 246 AA.
AC P00763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin II, anionic precursor (EC 3.4.21.4) (Pretirypsinogen II).
GN TRYP2.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
OX 11_TaxId=10116;
RP SEQUENCE FROM N.A.
RX MEDLINE=8505480; PubMed=6094547;
RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
RA Rutter W.J.;
RT "Structure of two related rat pancreatic trypsin genes."
RL J. Biol. Chem. 259:14255-14264 (1984).
RT [2]
RP SEQUENCE OF 9-246 FROM N.A.

```

RC STRAIN-Sprague-Dawley; TISSUE=Pancreas;
 RX MEDLINE=82245624; PubMed=6896710;
 RA McDonald R.J., Stary S.J., Swift G.H.;
 RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
 RL sequences of the cloned cDNAs." J. Biol. Chem. 257:9724-9732(1982).
 [3]
 RN X-RAY CRYSTALLOGRAPHY (1.59 ANGSTROMS).
 RP MEDLINE=93151936; PubMed=1381877;
 RA Earnest I., Faudon G., Craik C.S., Stroud R.J.
 RT "1.59-A structure of trypsin at 120 K: comparison of low temperature
 RL and room temperature structures." Proteins 10:171-187(1991).
 [4]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RP MEDLINE=96214506; PubMed=8634241;
 RA Brinen L.S., Willett M.S., Craik C.S., Fletcher R.J.;
 RT "X-ray structures of a designed binding site in trypsin show metal-
 RL dependent geometry." Biochemistry 35:5599-6009(1996).
 PL CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- SUBCELLULAR LOCATION: Extracellular.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL; V01274; CAA24581.1; -;
 DR EMBL; L00131; AAA98517.1; -;
 DR EMBL; L00130; AAA98517.1; JOINED.
 DR PIR; A22657; TRRT2.
 DR PDB; 1ANB; 01-APR-97.
 DR PDB; 1ANC; 01-APR-97.
 DR PDB; 1AND; 01-APR-97.
 DR PDB; 1ANE; 01-APR-97.
 DR PDB; 1AMH; 24-DEC-97.
 DR PDB; 1SLU; 11-JUL-96.
 DR PDB; 1SLV; 11-JUL-96.
 DR PDB; 1SLW; 11-JUL-96.
 DR PDB; 1SLX; 11-JUL-96.
 DR PDB; 1DPO; 07-JUL-97.
 DR PDB; 1TGI; 23-DEC-98.
 DR PDB; 1TGU; 23-DEC-98.
 DR MEROPS; S01.1258.1.1; Chymotrypsin.
 DR InterPro; IPR001254; Serine_protease_Try.
 DR Pfam; PF00408; trypsin.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_SPC; 1.
 DR PROSITE; PS00340; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family; 3D-structure.
 FT SIGNAL 1 15
 FT PROPEP 16 23 ACTIVATION PEPTIDE.
 FT CHAIN 24 246 TRYPSIN I1, ANIONIC.
 FT CHAIN 24 246 TRYPSIN I1, ANIONIC.
 FT ACT SITE 63 63 CHARGE RELAY SYSTEM.
 FT ACT SITE 107 107 CHARGE RELAY SYSTEM.
 FT ACT SITE 200 200 CHARGE RELAY SYSTEM.
 FT ACT SITE 200 200 CHARGE RELAY SYSTEM.
 FT DISULFID 30 160
 FT DISULFID 48 64
 FT DISULFID 132 233
 FT DISULFID 139 206
 FT DISULFID 171 185
 FT DISULFID 196 220
 FT DISULFID 194 194
 FT SITE 194 194
 FT CONFLICT 84 84

 CC REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 CC N -> D (IN REF. 1).

FT CONFLICT 88 88 V -> I (IN REF. 1).
 SQ SEQUENCE 246 AA; 26228 MW; ABD3630809AEB606 CRC64;
 Query Match 62.7%; Score 726; DB 1; Length 246;
 Best Local Similarity 61.6%; Pred. No. 2.1e-67;
 Matches 138; Conservative 25; Mismatches 57; Indels 4; Gaps 3;
 QY 1 IVGGVECTKHSQAHQVSLNSGCHKCGSLINXXVWVSAACHYKSVLVRGLGHHIRNVEG 60
 DB 24 IVGGVTCQNSVPQVSLNSGYPFCGSLINDQWVWVSAACHYKSRIVRGLGHHNINVEG 83
 QY 61 TRQXISSSVXRRHPNYSYNTXNDIMLKLPKXPATLQNYVNAVLPTFCADATCTCTUG 120
 DB 84 NEQFVNAAKIKHPNDFRKTNNNDIMLKLSVPKLNARVATVLPSCAPACTOCLISG 143
 QY 121 WGNVTSS-VVGDGDKLQKLPLFLSHADCAANSYPGMITQSMFCYGLGEGKSCQGDSSG 179
 DB 144 WGNLTSSGVNEPDLQCLDAPLPQADCEAST-PGKITDMVVCVFLGEGKSCQGDSSG 202
 QY 180 PWCNVQLQGVGVWNGVGYCABRXDYGKAVKXVLSGVGDTHA 223
 DB 203 PWCNGELQ-GIVSGTGCALDPNFGVTKVCNVYDVIQDTIA 244
 RESULT 12
 TRX2_BOVIN STANDARD; PRT; 247 AA.
 AC Q25463; 1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin, anionic precursor (EC 3.4.21.4).
 OS Eukaryota (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Holstein-Friesian; TISSUE=Pancreas;
 RA MEDLINE=91065383; PubMed=1701147;
 RX Le Huouou I., Wicker C., Guilloteau P., Toullec R., Puigserver A.;
 RT "Isolation and nucleotide sequence of cDNA clone for bovine
 RT pancreatic anionic trypsinogen. Structural identity within the
 RT trypsin family." Eur. J. Biochem. 193:767-773(1990).
 RL Eur. J. Biochem. 193:767-773(1990).
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -|- SUBCELLULAR LOCATION: Extracellular.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL; X54703; CAA38513.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.1258.1.1; Chymotrypsin.
 DR InterPro; IPR001314; Serine_protease_Try.
 DR Pfam; PF00089; Crpsaln; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS00340; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
 KW Multigene family; 3D-structure.
 FT SIGNAL 1 15
 FT PROPEP 16 23 ACTIVATION PEPTIDE.
 FT CHAIN 24 247 TRYPSIN, ANIONIC.

```

FT ACT SITE      63   63
CHARGE RELAY SYSTEM.
FT ACT SITE     107  107
CHARGE RELAY SYSTEM.
FT ACT SITE     200  200
CHARGE RELAY SYSTEM.
FT DISULFID     30  160
BY SIMILARITY.
FT DISULFID     48   64
BY SIMILARITY.
FT DISULFID    132  233
BY SIMILARITY.
FT DISULFID    139  206
BY SIMILARITY.
FT DISULFID    171  185
BY SIMILARITY.
FT DISULFID    196  220
BY SIMILARITY.
SQ SEQUENCE    247 AA; 26289 MM; 50A070495A7731DB CR664;
Query Match
Best Local Similarity 62.3%; Score 721; DB 1; Length 247;
Matches 136; Conservative 30; Mismatches 54; Indels 4; Gaps 3;
Oy 1 IVGVGYCTKHSHQAHVSLNGHVGXGGLLKKVWVSAAHYKSVLPRLGEHFHHITVNSG 60
Db 24 IVGVGYTCAENSVPVQVLSNAGYHFCGGSLINDQWVVSAHCYQHIVELGEFYNDVLGS 83
Oy 61 TEQMISSESVRRHPNVSSYNINXNDIMIKLTXTATLNQYHVAVALPTCEADANTACTVSG 120
Db 84 GRFOFDASKIRHPKYVSWTLNDILLIKLSTPAVINARVSTLLPSACASAGTECLISG 143
Oy 121 WGNWNS - VYGDKLOKLXPILSHADCANSGPMTQSNECAGYLEGGDCQCGDSG 179
Db 144 WGNLSSGWNTPDLCUAPLPLSHADCASY-FQGTTNMTCAGFLGGDCQCGDSG 202
Oy 180 PVVCNGVLGVGVWSKVGCAERDXRVYAKRVLSGWRDPM 223
Db 203 PVANCQLQ--GIVSWGYGCAKGKGVTVNYGVMDIETIA 244

RESULT 13
ID TRV3 RAT STANDARD; PRT: 247 AA.
AC P08426;
DT 01-AUG-1998 (Rel. 08, Created)
DE 01-AUG-1998 (Rel. 08, Last sequence update)
DT 15-JUN-2011 (Rel. 41, Last annotation update)
DE Trypsin III, cationic precursor (EC 3.4.21.4) (Pretrypsinogen III).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxId=10116;
RX [1]
RP SEQUENCE FROM N.A.
RF MEDLINE=87271609; PubMed=3607011;
RA Fletcher T.S., Alhadeff M., Craik C.S., Largman C.;
RT Isolation and characterization of a cDNA encoding rat cationic
RI Strypsinogen.; 6:3081-3086(1987).
RL - CHARGES: +6; 100% identical to extracellular.
CC - SUBCELLULAR LOCATION: Extracellular.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC This SWISS-PROT entry is copyright . It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is preserved in its
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/termsand
or send an email to license@ebi.ac.uk).
DR F0847; M15624; AAA1985.1; -.
DR P0847; A15647.
DR HSP8; P00763; AT5497.
DR MEROPS; S01.056f.
DR InterPro; IPR001314; Chymotrypsin.
DR DrPro; IPR001254; Ser protease_Try.
DR Pfam; PF00089; trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
```

```
CC or send an email to license@lsb-slb.ch).
CC -----
CC EMBL; X53458; CAA37538.1; -.
CC DR PIR; A35871; A35871.
CC DR HSSP; P00763; IDPO.
CC DR MEROPS; S01.151; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYP.
CC DR SMART; SM00020; TRYPSIN; 1.
CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KM Hydroxylase; Serine protease; Digestion; Zymogen; Signal;
CC Multigene family.
CC SIGNAL 1 15
CC FT PROPEP 16 20 ACTIVATION PEPTIDE.
CC FT CHAIN 16 20 TRYPSIN.
CC FT ACT_SITE 26 60 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 104 104 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 27 157 BY SIMILARITY.
CC FT DISULFID 45 61 BY SIMILARITY.
CC FT DISULFID 129 230 BY SIMILARITY.
CC FT DISULFID 136 203 BY SIMILARITY.
CC FT DISULFID 168 182 BY SIMILARITY.
CC FT DISULFID 193 217 BY SIMILARITY.
CC FT SITE 191 191 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
CC SQ SEQUENCE 243 AA; 25492 MW; C5B8345A8B3F8031 CRC64;

Query Match 62.0%; Score 718; DB 1; Length 243;
Best Local Similarity 61.2%; Pred. No. 1.4e-66;
Matches 137; Conservative 31; Mismatches 52; Indels 4; Gaps 3;

OY 1 IVGGVETXHSQAHQVSLNSGYKCGSLINXKXVWSAAHCYKSVLRVLGEHHRVNEG 60
DB 21 IIGGATCAKSPVIVVSLNSGYHFGGSLITNQVWSAAHCYKSIQVRLGEHHRVNEG 80
OY 61 TEQXISSSVVRHPNYSYINXNDIMLKITPATLQNYHAVALPTECADATMCTVSG 120
DB 81 TEQPISSKVRHSGVNSTYDNDIMLKSSPASLAAVNTVPLPGCSAGTSCILSG 140
OY 121 MNTMSVVD-GDKLQXLPILSHADANSYQGMITQSMFCAGYLEGGKDCSGGSG 179
DB 141 MNTLSNGSNVFDLLQCLNPLITNAQCNAY-PGEIYNNICVMEGGKDCSGGSG 199
OY 180 PWCNGVGVGVVSGVGHCHRDHXGKAYKXVLSCGWVDTMA 223
DB 200 PWCNGVGVGVVSGVGHCHRDHXGKAYKXVLSCGWVDTMA 241

RESULT 15
ID TR12_HUMAN STANDARD; PRT; 247 AA.
AC P07498 (Rel. 07, Created)
DT 01-APR-1998 (Rel. 07, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE T5-JUN-11 precursor (EC 3.4.21.4) (Anionic trypsinogen).
GN PRS2 OR TR2 OR TRYP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86221712; PubMed=3011602;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Mori T.,
RA Mateubara K.;
RT "Cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsinogens.";
RL Gene 41:305-310(1986).
```

```

RN [2]
RX MEDLINE=90091010; PubMed=2598466;
RA Kimland M., Russick C., Marks W.H., Borgstrom A.;
RT "Immunoreactive anionic and cationic trypsin in human serum.";
RL Clin. Chim. Acta 184:31-46(1989).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@lsb-slb.ch).
CC -----
CC EMBL; M27602; AAA61232.1; -.
CC DR PIR; B25852; B25852.
CC DR HSSP; P00763; IDPO.
CC DR MEROPS; S01.258; -.
CC DR Genew; HGNC:9483; PRS2.
CC DR MIN; 601564; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYP.
CC DR SMART; SM00020; TRYPSIN; 1.
CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KM Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
CC Multigene family.
CC SIGNAL 15 23 ACTIVATION PEPTIDE.
CC FT PROPEP 16 23
CC FT CHAIN 16 23 TRYPSIN II.
CC FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 30 160 BY SIMILARITY.
CC FT DISULFID 48 64 BY SIMILARITY.
CC FT DISULFID 171 185 BY SIMILARITY.
CC FT DISULFID 196 220 BY SIMILARITY.
CC FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
CC SQ SEQUENCE 247 AA; 26488 MW; 82B0F41EBB83D5DB CRC64;

Query Match 61.3%; Score 710; DB 1; Length 247;
Best Local Similarity 60.3%; Pred. No. 9.5e-66;
Matches 135; Conservative 31; Mismatches 54; Indels 4; Gaps 3;

OY 1 IVGGVETXHSQAHQVSLNSGYKCGSLINXKXVWSAAHCYKSVLRVLGEHHRVNEG 60
DB 24 IVGGVETXHSQAHQVSLNSGYKCGSLINXKXVWSAAHCYKSVLRVLGEHHRVNEG 83
OY 61 TEQXISSSVVRHPNYSYINXNDIMLKITPATLQNYHAVALPTECADATMCTVSG 120
DB 84 TEQPISSKVRHSGVNSTYDNDIMLKSSPASLAAVNTVPLPGCSAGTSCILSG 143
OY 121 MNTMSVVD-GDKLQXLPILSHADANSYQGMITQSMFCAGYLEGGKDCSGGSG 179
DB 144 MNTLSNGSNVFDLLQCLNPLITNAQCNAY-PGEIYNNICVMEGGKDCSGGSG 202
OY 180 PWCNGVGVGVVSGVGHCHRDHXGKAYKXVLSCGWVDTMA 223
DB 203 PWCNGVGVGVVSGVGHCHRDHXGKAYKXVLSCGWVDTMA 244

Search completed, February 12, 2003, 10:23:21
Job time : 31.2339 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 10:03:55 ; Search time 151.119 Seconds
(without alignments)
198.195 Million cell updates/sec

Title: US-10-036-371-8

Perfect score: 1158

Sequence: 1 IVGGYCTXSHQHQSLNS.....GVYAKVVLGSRDTHAY 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Watch 0s

Listing first 45 summaries

Database :

```

1: /SID52/gcgdata/geneexp-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneexp-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneexp-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneexp-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneexp-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneexp-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneexp-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneexp-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneexp-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneexp-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneexp-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneexp-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneexp-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneexp-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneexp-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneexp-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneexp-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneexp-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneexp-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneexp-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneexp-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneexp-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneexp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the posted score. The result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	1130	97.6	225	22 AAB31579 Amino acid sequenc
2	886	76.5	240	23 ABB04644 Engraulis japonicu
3	983	76.3	241	23 ABB04645 Engraulis japonicu
4	727	65.2	246	19 AAW44260 Human amyloid beta
5	727	62.9	247	21 AAY78974 Canine cationic tr
6	715	61.3	247	21 AAB35701 Human trypsin hL a
7	710	61.3	247	21 AAB35701 Human pancreatic t
8	707	61.1	247	18 AAG67693 Porcine pancreatic
9	706	61.0	246	23 AAY78975 Canine anionic try
10	702	60.6	223	23 ABB83322 Partial trypsin se

11	702	60.6	224	20 AAY31160 Human trypsin seri
12	702	60.6	225	22 AAB30816 Human trypsin seri
13	702	60.6	230	20 AAB93488 Human trypsinogen
14	702	60.6	247	21 AAB21321 Bovine trypsinogen
15	701	60.5	223	20 AAB81767 Bovine trypsinogen
16	701	60.5	223	21 AAY69973 Bovine trypsinogen
17	701	60.5	224	15 AAR53637 Bovine trypsinogen
18	701	60.5	229	21 AAY77494 Bovine trypsinogen
19	701	60.5	231	21 AAY91926 Recombinant trypsi
20	701	60.5	231	22 AAB60953 Trypsinogen analog
21	701	60.5	233	21 AAY91925 Trypsinogen-like p
22	700	60.5	230	15 AAR53638 Human trypsinogen
23	692	59.8	241	21 AAB21316 Human trypsinogen
24	692	59.7	240	19 AAB57740 Trypsinogen-like p
25	662	57.4	247	15 AAR53637 Human trypsinogen
26	665	57.4	247	15 AAR53637 Human trypsinogen
27	500	43.2	238	20 AAB60953 Human pancreatic t
28	494	42.7	256	23 AAY78974 Chimeric serine pr
29	494	42.7	320	23 AAY78974 Human protease, PR
30	494	42.7	320	23 AAY78974 Human protease, PR
31	466	40.2	151	22 AAU68927 Amino acid sequenc
32	460	39.7	217	21 AAB21305 Human trypsinogen
33	460	39.7	251	21 AAB21305 Human trypsinogen
34	451	38.9	250	21 AAB21325 Human trypsinogen
35	451	38.9	250	21 AAY99390 Human trypsinogen
36	451	38.9	250	21 AAB50479 Human trypsinogen
37	451	38.9	250	22 AAB50479 Human trypsinogen
38	451	38.9	250	22 AAB66139 Protein of the inv
39	451	38.9	250	23 AAB66139 Protein of the inv
40	451	38.9	250	23 AAB66139 Protein of the inv
41	451	38.9	250	23 AAB66139 Protein of the inv
42	451	38.9	250	23 AAB66139 Protein of the inv
43	451	38.9	250	23 AAB66139 Protein of the inv
44	451	38.9	250	23 AAB66139 Protein of the inv
45	451	38.9	250	23 AAB66139 Protein of the inv

ALIGNMENTS

RESULT 1

ID AAB31579 standard; peptide; 225 AA.

XX AAB31579;

XX AAB31579;

DT 20-APR-2001 (first entry)

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

XX AAB31579;

FT	/label= Ile, Xaa
FT	/note= "Xaa is any amino acid or absent"
Misc-difference 82	
FT	/label= Asp, Asn
FT	
Misc-difference 92	
FT	/label= Lys, Glu
FT	
Misc-difference 129	
FT	/label= Ala, Asp
FT	
Misc-difference 136	
FT	/label= Val, Cys
FT	
Misc-difference 138	
FT	/label= Asn, Ser
FT	
Misc-difference 204	
FT	/label= His, Asn
FT	
Misc-difference 212	
FT	/label= Met, Val, Cys
FT	
Misc-difference 224	
FT	/label= Asn, Ser
FT	
XX	WQ200078332-A2.
PD	
PD	28-DEC-2000.
XX	
XX	15-JUN-2000; 2000HO-IS000005.
XX	
PD	
PD	18-JUN-1999; 99IS-0005086.
XX	
PA	(BJAR/) BJARNASON J B.
PI	
PI	Bjarnason JB;
DR	
XX	WPI; 2001-091493/10.
XX	
Fish serine proteinase, useful as a cosmetic, medicament for treating eczema, psoriasis, arthritis, and in the manufacture of the medicament for treating, preventing pathogenic diseases involving receptor mediated binding -	
XX	
PS	Disclosure; Page 13; 38pp; English.
XX	
CC	The specification describes a fish serine proteinase. The proteinases are useful as medicaments, for treating and preventing a disease in a human or an animal such as pain, acute inflammation, chronic arthritis, inflammation, arthritis, inflamed joints, bursitis, osteoarthritis, rheumatoid arthritis, juvenile rheumatoid arthritis, septic arthritis, fibromyalgia, systemic lupus erythematosus, phlebitis tendinitis, rashes, psoriasis, acne, skin infections, seboreic eczema, eczema of the hands and feet, contact allergic reactions, athlete's foot, fungal infections, insect bites, ulcers, navel infections in newborns, wrinkles, scars, keloids, boils, warts and allergic itch, hemorrhoids, wounds, wound infections, wounds from burns, fungal infections and immunological and autoimmune diseases. They are also useful for removing dead or peeling skin from otherwise healthy skin, and for treating or preventing a disease in which pathogenesis is caused by bacteria, virus, fungus, parasite or a protozoan or a receptor mediated binding is involved. The present sequence represents the amino acid sequence of cod trypsin-like isozymes.
XX	
SQ	Sequence 225 AA;
Query Match	97.6%; Score 1130; DB 22; Length 225;
Best Local Similarity	100.0%; Pred. No. 7.2e-96;
Matches 225; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Oy	1 IVGGYECTXSQHQSLSNGHHVCGSLINXKVVSAAHCKYSVLRYLGRHHTFNEG 60
Dd	1 IVGGYECTXSQHQSLSNGHHVCGSLINXKVVSAAHCKYSVLRYLGRHHTFNEG 60
Oy	61 TEQTSSSSTQRHNSYNIDMLKLTPATLNQYHVVALPTFCADATCTVSG 120
Dd	61 TEQTSSSSTQRHNSYNIDMLKLTPATLNQYHVVALPTFCADATCTVSG 120
Oy	121 WQWWSVSVPDCKIQLVILPFLSHADCANVSGPMWTOSMEFGYLEGSKDS CGSDSGP 180

QY 121 MONTNSSYXDKQKQLXLPILSHADCANSGPGMITOSMFCAGYLEGGKDSQQSGGP 180
 DB 140 MONTMSYV-SGQKQLCQIQPILSDRCDSNSY-PGMITDMFCAGYLEGGKDSQQSGGP 197
 QY 181 VVNGVGLQGVVSWGSGYGCASRDXPQVYAKVYVLSGWVDTMA 223
 DB 198 VVNGELQ--GVWSGSGYGCASRDHPQVYAKVCIPTDMLQSTMA 238

RESULT 3

AB04645
 ID ABB04645 standard; Protein; 241 AA.

AC ABB04645;
 XX
 DT 04-MAR-2002 (first entry)
 XX

DE Engraulis japonicus trypsinogen (attr I1) SEQ ID NO:2.

XX Protease; trypsinogen; sardine; Japanese anchovy; fish sauce.

XX Engraulis japonicus.

XX JP2001269173-A.

XX 02-OCT-2001.

XX 24-MAR-2000; 2000JP-0084302.

XX 24-MAR-2000; 2000JP-0084302.

XX (NTSB) JAPAN TOBACCO INC.

XX MPI; 2002-078276/11.

XX N-PSDB; ABA04346.

XX A new DNA sequence -

XX Claim 1; Fig 8; 32pp; Japanese.

CC The present sequence represents a trypsinogen, which is a protease (I1) isolated from Engraulis japonicus (a fish called Engraulis japonica or Japanese anchovy). The present invention also relates to a DNA encoding (I1), or encoding a protease consisting of the amino acid sequence in which part of the amino acid residue is replaced inserted or deleted in the amino acid sequence encoded by the above DNA and having a bio-activity substantially same as (I1); (2) an expression vector in which the above DNA is recombined; (3) producing a sardine-derived protease in which a host cell transformed by the above expression vector is cultured and (I1) is recovered; (4) a protease containing substantially no other protein derived from fish; and (5) the preparation of a fish sauce in a short period while inhibiting the generation of an unpleasant smell compared to a case where the protease prepared by the above method is not added in which at least one of (1) to (5) is used. A shellfish selected from the group consisting of Clupeidae order is added to the aqueous solution containing salts in as high salt concentration as about 8% to 24% and the protease prepared by the above method is added to it. The method is used for about 1 to 11 months. The method is used for the preparation of a fish sauce in a short period.

XX Sequence 241 AA;

Query Match 76.3%; Score 883; DB 23; Length 241;
 Best Local Similarity 74.2%; Pred. No. 3.4e-73;
 Matches 167; Conservative 14; Mismatches 40; Indels 4; Gaps 3;

QY 1 IVGGVETCTHXSQAQVSLNSGVHXCXGSLINXXVWSAHCYKSVLVRIGEHHRVNEG 60

DB 20 IVGGVETCTPSPHQVSLNSGVHFCXGSLSDSWWSAHCYKSVLRVVRGHEHIMTEG 79

QY 61 TEQKISSSVXKHPNYSNNYNDIMLTKLTXPATLNQVTHVALPTECAADATNCTVSG 120

DB 80 NCOFIDSRVTRIPQYOTSYNDIMLTKLTKPATLNQVTHVALPTECAPAGTNCVLVG 139
 QY 121 MONTMSYVSGQKQLXLPILSHADCANSGPGMITOSMFCAGYLEGGKDSQQSGGP 180
 DB 140 MONTMSYV-SGQKQLCQIQPILSDRCDSNSY-PGMITDMFCAGYLEGGKDSQQSGGP 197
 QY 181 VVNGVGLQGVVSWGSGYGCASRDXPQVYAKVYVLSGWVDTMA 225
 DB 198 VVNGELQ--GVWSGSGYGCASRDHPQVYAKVCLFNWDISTMAQY 240

RESULT 4

AA064260
 ID AA064260 standard; Protein; 246 AA.

XX

XX AA064260;

XX 24-NOV-1998 (first entry)

XX Human amyloid beta-protein precursor inhibitor.

XX MCP-7; mast cell protease 7; tryptase-7; serine protease;

XX amyloid beta-protein precursor inhibitor; human blood clot;
 XX anticoagulant; myocardial infarction; reocclusion; thromboembolism;
 XX cerebral embolism; thrombosis; therapy.

XX Homo sapiens.

XX WO9824886-A1.

XX 11-JUN-1998.

XX 25-NOV-1997; 97WO-US21620.

XX 04-DEC-1996; 96US-0032354.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL.

XX Stevens RL;

XX WPI; 1998-333308/29.

XX New compositions containing tryptase-7, e.g. mouse mast cell
 PT protease-7 - are used to treat clot formation in e.g. myocardial
 PT infarction, reocclusion following angioplasty or pulmonary
 PT thromboembolism

XX Example; Page 77; 52pp; English.

CC This polypeptide comprises the human amyloid beta-protein precursor
 CC inhibitor (AβPI). The crystallographic structure of the complex of
 CC rat trypsin with AβPI was used to model the structure of the
 CC structure of the substrate-binding pocket of mouse mast cell
 CC protease 7 (mMCP-7, see AA064233). The invention relates to mMCP-7
 CC and related trypsin-like proteases that can be used to prevent or
 CC inhibit fibrin clot formation. Such proteases can be used to treat
 CC disorders mediated by undesirable thrombus clot formation such as
 CC myocardial infarction and reocclusion following angioplasty of
 CC blood clots associated with pulmonary thromboembolism, deep vein
 CC thrombosis, cerebral embolism, renal vein and peripheral arterial
 CC thrombosis.

XX Sequence 246 AA;

Query Match 65.2%; Score 755; DB 19; Length 246;
 Best Local Similarity 60.8%; Pred. No. 1.9e-61;
 Matches 143; Conservative 25; Mismatches 52; Indels 4; Gaps 3;

QY 1 IVGGVETCTHXSQAQVSLNSGVHXCXGSLINXXVWSAHCYKSVLVRIGEHHRVNEG 60

DB 24 IVGGVETCTPSPHQVSLNSGVHFCXGSLSDSWWSAHCYKSVLRVVRGHEHIMTEG 83

RESULT 7
 AAU87693
 ID AAU87693 standard; Protein; 247 AA.
 AC AAU87693;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human pancreatic tumour protein #5.
 XX
 KW Human; pancreatic tumour protein; immune response; pancreatic cancer;
 KW development of cancer; cancer progression; cytostatic.
 XX
 OS Homo sapiens.
 XX
 WN W0200012331-A2.
 XX
 PX 14-FEB-2002.
 XX
 PF 06-AUG-2001; 2001WO-US24619.
 XX
 PR 07-AUG-2000; 2000US-223130P.
 PR 30-JAN-2001; 2001US-265447P.
 PR 15-MAY-2001; 2001US-291201P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Pyle BA, Xu J, Kalos MD;
 XX
 DR WPI; 2002-241741/29.
 DR N-PSDB; ABK44208.
 XX
 PT Novel polynucleotide encoding pancreatic tumour polypeptides, useful in
 PT pharmaceutical compositions, e.g. vaccines, for treating pancreatic
 PT cancers.
 XX
 PS Claim 2; Page 164-165; 167pp; English.
 XX
 CC The present invention relates to the isolation of cDNA sequences
 CC (ABK44061-ABK44209) encoding human pancreatic tumour proteins. The
 CC polynucleotide sequences encoding human pancreatic tumour proteins are
 CC useful for stimulating an immune response in a patient and treating
 CC pancreatic cancer in a patient. A host cell that expresses these
 CC polynucleotides is useful for determining the presence of cancer in a
 CC patient. A composition comprising the polynucleotide, its encoded
 CC protein, or an antibody that binds to the protein may be used in the
 CC diagnosis, prevention and/or treatment of diseases, particularly
 CC pancreatic cancer. The sequences of the invention are also useful in
 CC pharmaceutical compositions, e.g. vaccines, for the diagnosis and
 CC inhibition of pancreatic cancer. Such compositions may be useful for
 CC inhibiting the development of cancer in a patient, or as markers for
 CC the progression of cancer. The polynucleotide sequences may also be used
 CC as probes or primers to nucleic acid hybridisation assays.
 CC AAU87689-AAU87694 represent human pancreatic tumour proteins.
 XX
 SQ Sequence 247 AA;
 Query Match 61.3%; Score 710; DB 23; Length 247;
 Best Local Similarity 60.3%; Pred. No. 2.6e-57;
 Matches 135; Conservative 31; Mismatches 54; Indels 4; Gaps 3;
 QY 1 IVGGYECTXHSQAHQVLSNGYHXCQGSILNXXWVSAACYSKSVLRVLGEHHRVNEG 60
 DB 24 IVGGYTCENSVFPVQVLSNGYHFCQGSILNXXWVSAACYSKSVLRVLGEHHRVNEG 83
 QY 61 TEQXISSSVYHKNYSYNYNDIMLKLYPATLQNYVHVALPTECAADATMCTVSG 120
 DB 84 NEQFINAAKITHPENFGNLTDLNLLKSSPATNSRVATVLSFRCAAGTECLISG 143
 QY 121 WGMTNSVVD-GDKLQXLLPLTSHADCANSGYGMWITOSMFCAGYLEGGKSCQGGSG 179
 DB 145 WGMTNSVVD-GDKLQXLLPLTSHADCANSGYGMWITOSMFCAGYLEGGKSCQGGSG 203

DB 144 WGMTLSGADYFPELQCLDAPVLSQAECEASY-PGKITNNMFCVGLGGKSCQGGSG 202
 QY 180 PVTNGVGLQGVNYSNGYHGCARDEXEYVYAKVLSQWVDTMA 223
 DB 203 PVTNSGELQ-GIVSWNGCAQKRGVITKTVYVDWIKOTIA 244
 RESULT 8
 AAU08475
 ID AAU08475 standard; Protein; 247 AA.
 XX
 AC AAU08475;
 XX
 DT 28-MAR-1997 (first entry)
 XX
 DE Porcine trypsinogen.
 XX
 KW Trypsinogen; trypsin; protease; filamentous fungus; Aspergillus.
 XX
 OS Sus scrofa.
 XX
 WN W09700316-A1.
 XX
 PX 03-JAN-1997.
 XX
 PF 10-JUN-1996; 96WO-DK00253.
 XX
 PR 16-JUN-1995; 95DK-0000693.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 DR Kjeldeen TB, Moldike HF;
 XX
 DR WPI; 1997-077523/07.
 DR N-PSDB; AAT49878.
 XX
 PT Prodn. of trypsin or trypsinogen in filamentous fungi - transformed
 PT with recombinant DNA comprising trypsinogen sequence fused to signal
 PT sequence, provides higher level of expression than other expression
 PT systems
 XX
 PS Example 2; Page 15; 28pp; English.
 XX
 CC An almost complete pig trypsinogen (AAU08475) lacking the very N-
 CC terminal end of the signal peptide is the product of a cDNA clone
 CC (AAT49878) isolated from a pig pancreas cDNA library. The cDNA clone
 CC can be fused to a signal sequence (e.g. from the native enzyme or
 CC a specified amylose gene) and used for prodn. of recombinant
 CC trypsinogen transformed host cells, esp. filamentous fungus, e.g.
 CC Aspergillus niger, Aspergillus niger, levels of trypsin
 CC secreted by Aspergillus niger are several-fold higher than those
 CC obtd. using other microbial systems.
 XX
 SQ Sequence 247 AA;
 Query Match 61.1%; Score 707; DB 18; Length 247;
 Best Local Similarity 60.3%; Pred. No. 4.8e-57;
 Matches 135; Conservative 27; Mismatches 56; Indels 4; Gaps 3;
 QY 1 IVGGYECTXHSQAHQVLSNGYHXCQGSILNXXWVSAACYSKSVLRVLGEHHRVNEG 60
 DB 25 IVGGYTCENSVFPVQVLSNGYHFCQGSILNXXWVSAACYSKSVLRVLGEHHRVNEG 84
 QY 61 TEQXISSSVYHKNYSYNYNDIMLKLYPATLQNYVHVALPTECAADATMCTVSG 120
 DB 85 NEQFINAAKITHPENFGNLTDLNLLKSSPATNSRVATVLSFRCAAGTECLISG 144
 QY 121 WGMTNSVVD-GDKLQXLLPLTSHADCANSGYGMWITOSMFCAGYLEGGKSCQGGSG 179
 DB 145 WGMTNSVVD-GDKLQXLLPLTSHADCANSGYGMWITOSMFCAGYLEGGKSCQGGSG 203
 QY 180 PVTNGVGLQGVNYSNGYHGCARDEXEYVYAKVLSQWVDTMA 223

DB 204 PWCNGQLQ--GIVSWGCAQKXKPGVTKVCNVNMIQTTA 245

RESULT 9
AAY78975
ID AAY78975 standard; protein; 246 AA.

XX AC AAY78975;
XX DT 05-JUN-2000 (first entry)
XX DE Canine anionic trypsinogen amino acid sequence.
XX KW Anionic trypsinogen; dog; monoclonal antibody production; detect;
XX KW diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
XX KW extrapancreatic hyposecretion.
XX OS Canis familiaris.
XX SN WO200009739-A1.
XX XX 24-FEB-2000.
XX PF 09-AUG-1999; 99WO-JP04299.
XX PR 10-AUG-1998; 98JP-0236609.
XX PR 10-MAR-1999; 99JP-0063990.
XX PA (FUJY) FUJI YAKUHIN KOGYO KK.
XX PI Waritani T, Ashida Y, Yamada T;
XX PI WPI; 2000-206019/18.
XX DR
XX PT Anti-canine trypsin monoclonal antibody, useful as reagent for quick
XX PT and accurate detection and quantitation of trypsin and/or trypsin-like
XX PT immunoreactants in various forms in diagnosis e.g. of pancreatic
XX PT diseases -
XX PS Claim 3; Page 64-65; 67pp; Japanese.

XX CC This sequence represents the canine anionic trypsinogen amino acid
XX CC sequence. The protein was isolated from the canine pancreas. The
XX CC invention relates to monoclonal antibodies with specificity against
XX CC canine trypsin, or canine trypsin-related substance. The antibodies are
XX CC highly specific and can be used as reagents for quick and accurate
XX CC detection and quantitation of canine trypsin and canine trypsin-like
XX CC immunoreactants in various forms. The antibodies can be used in the
XX CC diagnosis of diseases such as pancreatitis, pancreatic cancer, renal
XX CC insufficiency and extrapancreatic hyposecretion.

XX SQ Sequence 246 AA;
Query Match 61.08; Score 706; DB 21; Length 246;
Best Local Similarity 60.78; Pred. No. 5,9e-57;
Matches 136; Conservative 28; Mismatches 56; Indels 4; Gaps 3;

OY 1 IVGGVECTHSQAHQVSLNSGHHKCGSLINXAWVVSAAHCYKSVLRVLGHEHHRVNEG 60
DB 24 IVGGVTCRNSVPVQVSLNSGHHKCGSLINXAWVVSAAHCYKSVLRVLGHEHHRVNEG 83
OY 61 TEQXISSSVXVRHNPVSSYVNDIMLIKLTYPATLNOYVHAVALPTECAADATMCTVSG 120
DB 84 GEQFINAAKIIIRHPRVNTANDIMLIKLTSPATLNSRVSAIAPKSCPAAGTQCLISG 143
OY 121 WGNWTMSVXD-GDKLQKLXLPILSHADCANSGYFGMITOSMFCAGYLEGGKDSQCQDSGG 179
DB 144 WGNWTOSQNYPDVQLCLAPFLTSDVNCNAY-PGQISNNMCLGTMEGGKDSQCQDSGG 202
OY 180 PWCNGLQGVWSNGVYCAERDPXGVYAKVXVLSGWRDTMA 223
DB 203 PWCNGLQ--GIVSWGCAQKXKPGVTKVCNVNMIQTTA 244

RESULT 10
ABB83322
ID ABB83322 standard; protein; 223 AA.

XX AC ABB83322;
XX DT 30-AUG-2002 (first entry)
XX DE Partial trypsin sequence.
XX KW Trypsin; three dimensional structure; protein structure analysis;
XX KW serine protease.

XX OS Unidentified.
XX SN US2002035434-A1.
XX XX 21-MAR-2002.

XX PF 23-JUL-2001; 2001US-0910054.
XX PR 06-FEB-1992; 92JP-0021012.
XX PR 11-DEC-1992; 92JP-0331703.
XX PR 08-FEB-1993; 93US-0014867.
XX PA (FUIT) FUJITSU LTD.

XX PI Tomikawa M, Aikawa S, Matsuzawa F;
XX PI WPI; 2002-507172/54.

XX PT Analysis of three-dimensional structures by generating combination of
XX PT correspondence satisfying restriction condition, and calculating root
XX PT mean square distance between elements in the combination of
XX PT correspondence -
XX PS Disclosure; Fig 38; 65pp; English.

XX CC The present invention relates to a method for analysis of three
XX CC dimensional structures. The method involves generating a combination of
XX CC correspondence satisfying a restriction condition between the elements
XX CC belonging to a first and second point sets from among all candidates for
XX CC the combination of correspondence, and calculating a root mean square
XX CC distance between the elements corresponding in the combination of
XX CC correspondence. The method is useful for analysing three-dimensional
XX CC structures of molecules, particularly proteins. To illustrate the
XX CC invention, the serine proteases trypsin and elastase were used. The
XX CC present sequence is a partial sequence for trypsin. The histidine and
XX CC serine active sites of trypsin and elastase were analysed using the
XX CC method of the invention.

XX SQ Sequence 223 AA;
Query Match 60.64; Score 702; DB 23; Length 223;
Best Local Similarity 60.34; Pred. No. 1.2e-56;
Matches 135; Conservative 29; Mismatches 56; Indels 4; Gaps 3;

OY 1 IVGGVECTHSQAHQVSLNSGHHKCGSLINXAWVVSAAHCYKSVLRVLGHEHHRVNEG 60
DB 1 IVGGVTCRNTVPVQVSLNSGHHKCGSLINXAWVVSAAHCYKSVLRVLGHEHHRVNEG 60
OY 61 TEQXISSSVXVRHNPVSSYVNDIMLIKLTYPATLNOYVHAVALPTECAADATMCTVSG 120
DB 61 TEQVIFSKSVIVHPSVNSYNTLNDIMLIKLSAASLNSRVASISLPTSCASAGTQCLISG 120
OY 121 WGNWTMSVXD-GDKLQKLXLPILSHADCANSGYFGMITOSMFCAGYLEGGKDSQCQDSGG 179
DB 121 WGNWTSSGTYPDVQLCLAPFLTSDSSCKSAY-PGQITSNMFCAGYLEGGKDSQCQDSGG 179
OY 180 PWCNGLQGVWSNGVYCAERDPXGVYAKVXVLSGWRDTMA 223
DB 180 PWCNGLQ--GIVSWGCAQKXKPGVTKVCNVNMIQTTA 221


```
AAW81767
ID AAW81767 standard; peptide; 223 AA.
AC
XX AAW81767;
XX
DT 29-JAN-1999 (first entry)
XX
DE Bovine TRYP peptide fragment.
XX
KW DHFR; dihydrofolic acid reductase; protein function; trypsin; bovine;
KW ribonuclease; myoglobin; database; homology; resemblance.
XX
OS Bos taurus.
XX
PN JF10287696-A.
XX
XD 27-OCT-1998.
XX
PF 11-APR-1997; 97JP-0093577.
XX
PR 11-APR-1997; 97JP-0093577.
XX
PA (IYAK-) IYAKU BUNSHI SEKKEI KENKYUSHO KK.
XX
DR WPI; 1999-018384/02.
XX
CC Determination of protein biological function - comprises use of
PT amino acid sequences database containing the relevant information
XX
PS Example 1; Fig 1; 11pp; Japanese.
XX
CC This sequence is used in the creation of a database containing the
CC information for amino acid sequence of protein with at least 1
CC biological function with added a score on importance of expression of
CC the biological information for each amino acid residue. The database is
CC used for the determination of unknown biological function of a protein or
CC polypeptide. Based on the homology of amino acid sequence, e.g. steric
CC structure of the protein, the database includes retrieval and evaluation of high
CC homologous relationship for the determination of mostly resembling
CC protein. The database allows for the correct and rapid retrieval and
CC presumption of protein and polypeptide having biological functions.
XX
SQ Sequence 223 AA;

Query Match 60.5%; Score 701; DB 20; Length 223;
Best Local Similarity 60.3%; Pred. No. 1.5e-56;
Matches 135; Conservative 29; Mismatches 56; Indels 4; Gaps 3;

QY 1 IVGGYECTYHQAHOVLSNGYHXCGLSGLNXXWVSAACHYKSVLRVLGRHHIRVNEG 60
D6 1 IVGGTGCANTVPYQVLSNGYHFCGGLSGLNXXWVSAACHYKSGIOLRGEDINWVEG 60
QY 61 TEQKISSSVYRHNYSYNYNDIMLKITPATLNQYVAVALPTCAADATMCTVSG 120
D6 61 NEQFTSASKSVHFSYKSNLTNNDIMLKITXLSASLNSRVASISLPTSCASAGTQCLISG 120
QY 121 WGMTMSS-VXDGKLOKXLPLTSLHADCANSGRCMTOSWFCAYLEGCKSCQDSGG 179
D6 121 WGMTKSGTSPVOLKCLKAPILSDSSCKSAY-PQITSNMFCAYLEGCKSCQDSGG 179
QY 180 PIVCNGLQGVGVSWGVGCAERDXPGVIKXVLSQWRDTWA 223
D6 180 PIVCSGKLO--GIVSWGSGCAQKNKPGVYTKVCNVYSIKQTIA 221

Search completed: February 12, 2003, 10:22:29
Job time : 152.119 secs
```

THIS PAGE BLANK (USPTO)